

Tue Mar 2 09:56:36 2004

us-09-455-486-6.rag

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using SW model

Run on: March 1, 2004, 19:53:00 ; Search time 96 Seconds

(without alignments) 1336.214 Million cell updates/sec

Title: US-09-455-486-6

2351

Perfect score: 1 MBSISMNGSPKSLSEKLPN.....ALVLSIVLPDLQKCPD 454

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database: A. Geneseq\_29Jan04:\*

1: geneseqp1980s:.\*  
2: geneseqp1990s:.\*  
3: geneseqp2000s:.\*  
4: geneseqp2001s:.\*  
5: geneseqp2002s:.\*  
6: geneseqp2003as:.\*  
7: geneseqp2003bs:.\*  
8: geneseqp2004s:.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2351	100.0	454	AAE02781	Aae02781 Human six
2	2351	100.0	454	AAU01018	Aau01018 Human ORF
3	2351	100.0	454	ABG61933	Abg61933 Prostate
4	2351	100.0	454	AAU76538	Aau76538 Tumour-as
5	2351	100.0	454	AAU80190	Aau80190 Human PUM
6	2351	100.0	454	ABU08893	Abu08893 Tumour-pro
7	2351	99.9	454	ABU08893	Abu08893 Human pro
8	2351	97.6	450	AAE28951	Aae28951 Human STE
9	2351	97.6	450	ABU63312	Abu63312 Human tra
10	2351	97.4	490	AAU01018	Aau01018 Human six
11	2036	86.6	419	AAU01018	Aau01018 Human ORF
12	1972.5	83.9	456	ABG12306	Abg12306 Novel hum
13	1972.5	83.9	456	ABG00113	Abg00113 Novel hum
14	1738	73.9	1273	AAE02841	Aae02841 Human STE
15	1364.5	58.0	1082	ABP62883	Abp62883 Human pol
16	1316	55.0	488	AAAB49481	Aaab49481 Rat p-HYD
17	1316	55.0	526	ABR83365	Abbr83365 Murine tu
18	1272	54.1	488	ABR83365	Abbr83365 Human dtu
19	1269	54.0	488	ABR83365	Abbr83365 Human tum
20	1269	54.0	488	AAU04564	Aau04564 Human pro
21	1269	54.0	488	AAU04564	Aau04564 Human G-P
22	1269	54.0	488	AAU10220	Aau10220 Human six
23	1269	54.0	488	ABU60864	Abu60864 Human G-P
24	1245.5	53.0	487	AAAB49481	Aaab49481 Human P-H
25	1088	46.3	459	AAU74715	Aau74715 Human mem

26	1087	46.2	459	AAE02782	Aae02782 Human six
27	1087	46.2	459	ABP64820	Abp64820 Human pro
28	1085	46.2	459	AAU01018	Aau01018 Human six
29	1082	46.0	458	AAE02636	Aae02636 Human STE
30	1074.5	45.7	456	ABR49482	Abrr49482 Human G-H
31	901	38.3	173	AAV68195	Aav68195 Human STR
32	844	35.9	179	AAU18006	Aau18006 Peptide #
33	844	35.9	179	ABR37041	Abrr37041 Peptide #
34	844	35.9	179	AAU05017	Aau05017 Peptide #
35	844	35.9	179	ABR31810	Abrr31810 Peptide #
36	844	35.9	179	ABR22356	Abrr22356 Protein #
37	844	35.9	179	AAU70183	Aau70183 Human bon
38	844	35.9	179	AAU57768	Aau57768 Human bra
39	844	35.9	179	ABG51883	Abg51883 Human liv
40	844	35.9	179	AAU05646	Aau05646 Peptide #
41	844	35.9	179	ABG39817	Abg39817 Human pep
42	768	32.7	237	AAU04565	Aau04565 Human G-P
43	768	32.7	237	ABU60865	Abu60865 Human G-P
44	736	31.3	141	AAU52589	Aau52589 Human pro
45	717	30.5	267	ABU60886	Abu60886 Human G-P

## ALIGNMENTS

RESULT 1  
ID AAE02781 standard; protein: 454 AA.

XX AAE02781;  
XX 06-AUG-2001 (first entry)

DE Human six transmembrane epithelial antigen of prostate (STEAP)-2 protein.

XX Human, cytosolic; antiproliferative; vaccine; gene therapy;

KW six transmembrane epithelial antigen of the prostate-2; STEAP-2;

KW chromosome 7q21; cancer; prostate; colon; bladder; lung; ovarian;

XX pancreatic.

XX Homo sapiens.

OS Homo sapiens.

XX Homo sapiens.

XX Homo sapiens.

XX Homo sapiens.

XX Homo sapiens.

XX Homo sapiens.

XX Homo sapiens.

XX Homo sapiens.

XX Homo sapiens.

XX Homo sapiens.

XX Homo sapiens.

XX Homo sapiens.

XX Homo sapiens.

XX Homo sapiens.

XX Homo sapiens.

XX Homo sapiens.

XX Homo sapiens.

XX 06-DEC-2000; 2000WO-US03040.  
 XX 06-DEC-1999; 99US-00455486.  
 XX (UROG-) UROGENESYS INC.  
 PA Afar DEH, Hubert RS, Raitano AB, Saffran DC, Mitchell SC,  
 PI Paris M, Jakobovits A,  
 XX WPI; 2001-367804/38.  
 DR N-PSDB; AAD07072.  
 XX  
 PT New STEAP (six transmembrane epithelial antigen of the prostate)  
 PT proteins, expressed in human cancers, useful for detecting and treating  
 PT cancer.  
 XX  
 PS Claim 1; Fig 9A-9C; 187pp; English.  
 XX  
 CC The present sequence is human six transmembrane epithelial antigen of the  
 CC prostate (STEAP)-2 protein. STEAP-2 is a member of cell surface serpentine  
 CC transmembrane antigens. STEAP-2 gene is located on chromosome 7q21 and is  
 CC used in gene therapy. Inhibiting the development or progression of a  
 CC cancer (eg. prostate, colon, bladder, lung, ovarian and pancreatic)  
 CC expressing STEAP or inhibiting growth or killing cells expressing STEAP  
 CC in a patient, comprises administering a vaccine composition to the  
 CC patient. Treating a patient with a cancer that expresses STEAP, or  
 CC inhibiting growth or killing cells expressing STEAP, comprises  
 CC administering to the patient a vector encoding single chain monoclonal  
 CC antibody that comprises the variable domains of the heavy and light  
 CC chains of the monoclonal antibody that specifically binds to STEAP, such  
 CC that the vector delivers the single chain monoclonal antibody coding  
 CC sequence to the cancer cells and the encoded single chain monoclonal  
 CC antibody is expressed intracellularly. Note: This sequence is stated to  
 CC be the same as that being shown as SEQ ID NO:8 (AA02841) in sequence  
 CC listing of the specification. However both the sequences differ at  
 CC several positions  
 CC  
 XX Sequence 454 AA;  
 SQ  
 Query Match 100.0%; Score 2351; DB 4; Length 454;  
 Best Local Similarity 100.0%; Pred. No. 5e-245;  
 Matches 454; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MESISMGSPKSLSETCLPENGINGIKDARKVTGVIGSGDFAKSLTIRLCGYHYVIGS 60  
 DB 1 MESISMGSPKSLSETCLPENGINGIKDARKVTGVIGSGDFAKSLTIRLCGYHYVIGS 60  
 QY 61 RNPFASEFPFHVVDVTHEDALTKNIIFAIHREHYTSLMDRLHVLGKILIDVSNM 120  
 DB 61 RNPFASEFPFHVVDVTHEDALTKNIIFAIHREHYTSLMDRLHVLGKILIDVSNM 120  
 QY 121 RINPESNAEYLAISLPDSLIYGFVVSAMALQGPCKASQVYICSNICARQOIVE 180  
 DB 121 RINPESNAEYLAISLPDSLIYGFVVSAMALQGPCKASQVYICSNICARQOIVE 180  
 QY 181 LARQINFIPIDLGSLSSAREIENLPRLFTLMRGPVVAISLATPFPLVSFVNDVTHPYA 240  
 DB 181 LARQINFIPIDLGSLSSAREIENLPRLFTLMRGPVVAISLATPFPLVSFVNDVTHPYA 240  
 QY 241 RNOQSDPYKPIPIEVNKTLPPIVAITLISLVYLAGLAAAOVYGRKRPFPMTLWIO 300  
 DB 241 RNOQSDPYKPIPIEVNKTLPPIVAITLISLVYLAGLAAAOVYGRKRPFPMTLWIO 300  
 QY 301 CRKQLGLISFPFAMVAVASLCIPMRSEERYLFLLNMAVQOVHANINENNEEVRRIEYV 360  
 DB 301 CRKQLGLISFPFAMVAVASLCIPMRSEERYLFLLNMAVQOVHANINENNEEVRRIEYV 360  
 QY 361 ISFQIMSLGLSLAIVSIPSVSNALNWRBESFQSTLGVALLISTFHLIYGMKRAFE 420  
 DB 361 ISFQIMSLGLSLAIVSIPSVSNALNWRBESFQSTLGVALLISTFHLIYGMKRAFE 420  
 QY 421 EBYRYFTPNFVLAIVLPISIVIDLQLCRYPD 454

DB 421 EBYRYFTPNFVLAIVLPISIVIDLQLCRYPD 454  
 RESULT 2  
 AAU10188  
 ID AAU10188 standard; protein; 454 AA.  
 XX  
 XX AAU10188;  
 AC  
 XX  
 XX 16-JAN-2002 (first entry)  
 DT  
 XX  
 DE Human ORF2 of Six-Transmembrane Protein of Prostate 1.  
 XX  
 XX Human, Six-Transmembrane Protein of Prostate 1; STMP1; prostate cancer;  
 XX benign prostatic hyperplasia; acute prostatitis; testicular cancer;  
 XX cryptorchidism; testicular disorder; proliferative disorder; lymphoma;  
 XX leukemia; melanoma; ovarian cancer; breast cancer; pancreatic cancer;  
 XX liver cancer; lung cancer; cytostatic; ORF2.  
 XX  
 OS Homo sapiens.  
 XX  
 XX WO200172962-A2.  
 PN  
 XX  
 XX 04-OCT-2001.  
 PD  
 XX  
 XX 23-MAR-2001; 2001WO-US009410.  
 PF  
 XX  
 XX 24-MAR-2000; 2000US-0191929P.  
 PR  
 XX  
 XX (SAAT/) SAATCIOGU F.  
 PA  
 XX  
 XX Saaticoglu F;  
 PI  
 XX  
 XX WPI; 2001-662926/76.  
 DR  
 XX  
 XX N-PSDB; AAS15810; AAS15811.  
 DR  
 XX  
 PT New polynucleotide for the diagnosis, prevention and treatment for  
 PT prostate and testis disorders, particularly prostate cancer, comprises  
 PT prostate-specific or testis-specific nucleic acids.  
 PT  
 XX  
 XX Claim 1; Fig 4H; 11pp; English.  
 PS  
 XX  
 CC The invention relates to substantially pure prostate-specific or testis-  
 CC specific polypeptides and the nucleic acids encoding them. Also included  
 CC are vectors and host cells expressing the proteins, a transgenic animal  
 CC expressing the protein, antibodies against the proteins, probes for  
 CC detecting the nucleic acids, antisense molecules for the nucleic acids  
 CC and methods of isolating modulators of the proteins. Compounds that  
 CC modulate the prostate specific or testis specific polypeptide are useful  
 CC to diagnose, prevent or treat disorders of the testis or prostate  
 CC particularly prostate cancer, benign prostatic hyperplasia, acute  
 CC prostatitis, testicular cancer, cryptorchidism, undescended, retractile,  
 CC ascending or vanished testis. Other proliferative disorders for which the  
 CC modulators may be used include lymphoma, leukemia, melanoma, ovarian  
 CC cancer, breast cancer, pancreatic cancer, liver cancer and lung cancer.  
 CC The present sequence is prostate specific protein, Six-Transmembrane  
 CC Protein of Prostate 1, STMP1, ORF2  
 CC  
 XX Sequence 454 AA;  
 SQ  
 Query Match 100.0%; Score 2351; DB 4; Length 454;  
 Best Local Similarity 100.0%; Pred. No. 5e-245;  
 Matches 454; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MESISMGSPKSLSETCLPENGINGIKDARKVTGVIGSGDFAKSLTIRLCGYHYVIGS 60  
 DB 1 MESISMGSPKSLSETCLPENGINGIKDARKVTGVIGSGDFAKSLTIRLCGYHYVIGS 60  
 QY 61 RNPFASEFPFHVVDVTHEDALTKNIIFAIHREHYTSLMDRLHVLGKILIDVSNM 120  
 DB 61 RNPFASEFPFHVVDVTHEDALTKNIIFAIHREHYTSLMDRLHVLGKILIDVSNM 120

QY 121 RINQYPSNAEYLAIFPDLSLVKGNVVSAMALQGPDKASROYVICSNNIOARQOYIE 180  
 DB 121 RINQYPSNAEYLAIFPDLSLVKGNVVSAMALQGPDKASROYVICSNNIOARQOYIE 180  
 QY 181 IARQNLFIPIIDIGLSLSAREIENLPLRLFTLRGPNVVAISLATFFLYSFVRDVIHPYA 240  
 DB 181 IARQNLFIPIIDIGLSLSAREIENLPLRLFTLRGPNVVAISLATFFLYSFVRDVIHPYA 240  
 QY 241 RNOQSDPFYKIPRIEIVNKTPIVAITLLSLVYLAGLLAAAYQLYGTRRRPPELMTWQ 300  
 DB 241 RNOQSDPFYKIPRIEIVNKTPIVAITLLSLVYLAGLLAAAYQLYGTRRRPPELMTWQ 300  
 QY 301 CRKQGLISFPFAMVAVASLCLPMRBSERYFLNMAVQOVHANINENSNNEBEVRIEM 360  
 DB 301 CRKQGLISFPFAMVAVASLCLPMRBSERYFLNMAVQOVHANINENSNNEBEVRIEM 360  
 QY 361 ISFGIMSLGLSLAVTSIPSVSNALNWEFSFIQSTIGYVALLISTFHVLIYGMKRAFE 420  
 DB 361 ISFGIMSLGLSLAVTSIPSVSNALNWEFSFIQSTIGYVALLISTFHVLIYGMKRAFE 420  
 QY 421 EBYRFPYTPNPFVLAIVPSIYVILDLQCRYPD 454  
 DB 421 EBYRFPYTPNPFVLAIVPSIYVILDLQCRYPD 454

## RESULT 3

ABG61933  
 ID ABG61933 standard; protein; 454 AA.

AC ABG61933;

DT 15-AUG-2002 (first entry)

DE Prostate cancer-associated protein #134.

KM Prostate cancer; prostate tumour tissue; human; mammal; cytostatic.

OS Mammalia.

PN WO200230268-A2.

PD 18-APR-2002.

PF 12-OCT-2001; 2001WO-US032045.

PR 13-OCT-2000; 2000US-00687576.

PR 08-DEC-2000; 2000US-00733288.

PR 24-JAN-2001; 2001US-02639577.

PR 16-MAR-2001; 2001US-0276794P.

PR 06-MAR-2001; 2001US-0276888P.

PR 24-APR-2001; 2001US-0286214P.

PR 30-APR-2001; 2001US-00847046.

PR 04-MAY-2001; 2001US-0288589P.

PA (EOSB-) EOS BIOTECHNOLOGY INC.

PI Gish KC, Mack DH, Wilson KE, Afar D, Hevezi P;

DR WPI; 2002-471335/50.

PS N-PsDB; ABR62252.

PT Detecting a prostate cancer-associated transcript in a cell in a patient,

CC useful for diagnosing prostate cancer (PC) or screening modulators of PC, by determining if prostate cancer-associated genes are expressed in a prostate tissue.

Claim 27; Page 416; 436pp; English.

CC The present invention relates to methods of detecting a prostate cancer-associated transcript in a cell from a patient. The method comprises contacting a biological sample from the patient with prostate cancer-

CC associated polynucleotides (designated PC genes) that selectively hybridise to a sequence that is at least 80% identical to them. The prostate cancer-associated polynucleotide sequences are differentially expressed in prostate tumour tissue or in prostate cancer and are derived from the tissues of various organisms such as humans or other mammals (e.g. mice, sheep and dogs). The methods of the invention are useful for diagnosing and treating prostate cancer in mammals. The prostate cancer-associated genes are useful for diagnosing or treating prostate cancer, as well as for identifying modulators of prostate cancer or agents that inhibit prostate cancer. The nucleic acid sequences are particularly useful in gene therapy, as a vaccine or in antisense applications.

CC ABG61800-ABG61944 represent prostate cancer-associated proteins

SO Sequence 454 AA;

Query Match 100.0%; Score 2351; DB 5; Length 454;

Best Local Similarity 100.0%; Pred. No. 56-245;

Matches 454; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MESISMWGSFKSLSEFCIPGNGINGIDARKVTVGVSGDFAKSLTRLRGCHVTVGS 60

DB 1 MESISMWGSFKSLSEFCIPGNGINGIDARKVTVGVSGDFAKSLTRLRGCHVTVGS 60

QY 61 RNPFASEFPFHVVDVTHHEDALTKNIIIFVAIHREHYTSLMDLRHLVSKLIDIVSNM 120

DB 61 RNPFASEFPFHVVDVTHHEDALTKNIIIFVAIHREHYTSLMDLRHLVSKLIDIVSNM 120

QY 121 RINQYPSNAEYLAIFPDLSLVKGNVVSAMALQGPDKASROYVICSNNIOARQOYIE 180

DB 121 RINQYPSNAEYLAIFPDLSLVKGNVVSAMALQGPDKASROYVICSNNIOARQOYIE 180

QY 181 IARQNLFIPIIDIGLSLSAREIENLPLRLFTLRGPNVVAISLATFFLYSFVRDVIHPYA 240

DB 181 IARQNLFIPIIDIGLSLSAREIENLPLRLFTLRGPNVVAISLATFFLYSFVRDVIHPYA 240

QY 241 RNOQSDPFYKIPRIEIVNKTPIVAITLLSLVYLAGLLAAAYQLYGTRRRPPELMTWQ 300

DB 241 RNOQSDPFYKIPRIEIVNKTPIVAITLLSLVYLAGLLAAAYQLYGTRRRPPELMTWQ 300

QY 301 CRKQGLISFPFAMVAVASLCLPMRBSERYFLNMAVQOVHANINENSNNEBEVRIEM 360

DB 301 CRKQGLISFPFAMVAVASLCLPMRBSERYFLNMAVQOVHANINENSNNEBEVRIEM 360

QY 361 ISFGIMSLGLSLAVTSIPSVSNALNWEFSFIQSTIGYVALLISTFHVLIYGMKRAFE 420

DB 361 ISFGIMSLGLSLAVTSIPSVSNALNWEFSFIQSTIGYVALLISTFHVLIYGMKRAFE 420

QY 421 EBYRFPYTPNPFVLAIVPSIYVILDLQCRYPD 454

DB 421 EBYRFPYTPNPFVLAIVPSIYVILDLQCRYPD 454

## RESULT 4

AAU76538  
 ID AAU76538 standard; protein; 454 AA.

AC AAU76538;

DT 05-JUN-2002 (first entry)

DE Tumour-associated antigenic target protein, TAT138.

OS Homo sapiens.

Key Location/Qualifiers

FT Modified-site 8..13

FT Modified-site /note="N-myristoylation site"

FT Modified-site 24..29

FT /note= "N-myristoylation site"  
 FT Region 29..32  
 FT /note= "CAMP- and GMP-dependent protein kinase  
 FT phosphorylation site"  
 FT Modified-site 34..39  
 FT /note= "N-myristoylation site"  
 FT Modified-site 193..198  
 FT /note= "N-myristoylation site"  
 FT Domain 210..230  
 FT /note= "Transmembrane domain"  
 FT Modified-site 256..259  
 FT /note= "Aen is N-glycosylated"  
 FT Domain 257..277  
 FT /note= "Transmembrane domain"  
 FT Domain 259..379  
 FT /note= "Transmembrane domain"  
 FT Modified-site 274..279  
 FT /note= "N-myristoylation site"  
 FT Domain 299..319  
 FT /note= "Transmembrane domain"  
 FT Domain 393..413  
 FT /note= "Transmembrane domain"  
 FT Modified-site 415..424  
 FT /note= "Tyrosine kinase phosphorylation site"  
 FT Domain 428..448  
 FT /note= "Transmembrane domain"  
 PN WO200216429-A2.  
 PD 26-FEB-2002.  
 XX 22-JUN-2001; 2001WO-US020118.  
 XX 24-AUG-2000; 2000WO-US023328.  
 PR 26-SEP-2000; 2000US-0235451P.  
 PR 01-DEC-2000; 2000WO-US032678.  
 PR 28-FEB-2001; 2001WO-US006520.  
 PR 01-MAR-2001; 2001WO-US006666.  
 XX (GETH ) GENENTECH INC.  
 PA Goddard A, Godowski PJ, Gurney AL, Hillan KJ, Polakis P, Smith V;  
 PI Wood WI, Wu TD, Zhang Z,  
 XX N-PSDs; ABK11093.  
 DR Novel isolated tumor-associated antigenic target polypeptides which are  
 XX useful as targets for cancer therapy and diagnosis in mammals.  
 PT Claim 12; Fig 10; 121pp; English.  
 PS  
 CC The invention relates to an isolated tumour-associated antigenic target  
 CC polypeptide (TAT) (I), specifically TAT113-TAT138 polypeptides, and the  
 CC polynucleotides (II) encoding them. (II) is useful for diagnosing the  
 CC presence of a tumour in a mammal, where the level of expression of (II)  
 CC is indicative of the presence of tumour in the mammal from which the test  
 CC sample was obtained. Antibody to (I) is useful for killing a cancer cell  
 CC (e.g. breast cancer cell, a colorectal cancer cell, a lung cancer cell,  
 CC an ovarian cancer cell, a central nervous system (CNS) cancer cell, a  
 CC liver cancer cell, a bladder cancer cell, a pancreatic cancer cell, a  
 CC melanoma cell or a leukaemia cell) that expresses (I). Oligonucleotides  
 CC hybridising to (II) are useful as diagnostic probes, antisense  
 CC oligonucleotide probes or for encoding fragments of full length TAT  
 CC polypeptide. (II) is also useful in chromosome and gene mapping and in  
 CC the generation of antisense RNA and DNA probes, for constructing  
 CC hybridisation probes for mapping the gene encoding TAT and for genetic  
 CC analysis of individuals with genetic disorders. (II) is also useful for  
 CC generating either transgenic animals or knockout animals, and in gene  
 CC therapy. The TAT polypeptides and nucleic acids may also be used for  
 CC tissue typing and the TAT polypeptides are useful for screening compounds  
 CC that mimic the TAT polypeptide (agonists) or prevent the effect of TAT  
 CC polypeptide (antagonists). The antibody is useful for staging TAT

CC polypeptide-expressing cancers, purifying or immunoprecipitating TAT  
 CC polypeptide from cells for detection and quantitation of TAT polypeptide  
 CC in vitro, e.g., in an enzyme linked immunosorbent assay (ELISA) or  
 CC Western blot. The antibodies are also useful for treating a TAT-  
 CC expressing cancer or alleviating one or more symptoms of cancer in a  
 CC mammal. The present sequence represents the amino acid sequence of TAT1138  
 XX  
 SQ Sequence 454 AA:  
 Query Match 100.0%; Score 2351; DB 5; Length 454;  
 Best Local Similarity 100.0%; Pred. No. 5e-245;  
 Matches 454; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 MESISMGSPKSLSETCIPNGINGIKDARKYGVIGSGFASLTIRLRGCHVAVIGS 60  
 DB 1 MESISMGSPKSLSETCIPNGINGIKDARKYGVIGSGFASLTIRLRGCHVAVIGS 60  
 QY 61 RNPFRASEFPFHVVDVTHMEDALTKTNIIFVAIHREHYTSLMDLRHLVGKILIDVSNM 120  
 DB 61 RNPFRASEFPFHVVDVTHMEDALTKTNIIFVAIHREHYTSLMDLRHLVGKILIDVSNM 120  
 QY 121 RINGYPSNABEYLASLPDLSIVKGFVVSAMALQSPKASROYVCSNNIQARQVIE 180  
 DB 121 RINGYPSNABEYLASLPDLSIVKGFVVSAMALQSPKASROYVCSNNIQARQVIE 180  
 QY 181 LARQINFPIDIGLSSAREIENLPLRLFTLMRGPVVAISLATFPELYSFVRDVIHPYA 240  
 DB 181 LARQINFPIDIGLSSAREIENLPLRLFTLMRGPVVAISLATFPELYSFVRDVIHPYA 240  
 QY 241 RNOQSDPKPIEIVNKTLPVATTLISLYIAGLAAAOQYGYGTRRPPPLETWLQ 300  
 DB 241 RNOQSDPKPIEIVNKTLPVATTLISLYIAGLAAAOQYGYGTRRPPPLETWLQ 300  
 QY 301 CRKQGLISFPFANVHAAYSICLPMRSERYLFLNMAVQOVHANIENSNMEEVWRIEMY 360  
 DB 301 CRKQGLISFPFANVHAAYSICLPMRSERYLFLNMAVQOVHANIENSNMEEVWRIEMY 360  
 QY 361 ISFGIMSLGLISLAIVTSIPSVSNALNWRBSPFQSTLGYVALLISTFHVLIYWKRAFE 420  
 DB 361 ISFGIMSLGLISLAIVTSIPSVSNALNWRBSPFQSTLGYVALLISTFHVLIYWKRAFE 420  
 QY 421 EBYRFTYPPNFVALVLPSTVILDLQCRYPD 454  
 DB 421 EBYRFTYPPNFVALVLPSTVILDLQCRYPD 454  
 RESULT 5  
 AAU80190  
 ID AAU80190 standard; protein; 454 AA.  
 XX  
 AC AAU80190;  
 XX  
 DT 15-JUL-2002 (first entry)  
 XX  
 DE Human PUMPCn protein, PRO23203.  
 XX  
 KW Human; protein upregulated in metastatic prostate cancer; immunogen;  
 XX PUMPCn; PRO23203; prostate cancer; cytostatic; gene therapy; transgenic;  
 XX androgen independent prostate cancer; DNA microarray.  
 OS Homo sapiens.  
 XX  
 PN WO200226822-A2.  
 XX  
 PD 04-APR-2002.  
 XX  
 PF 26-SEP-2001; 2001WO-US030290.  
 PR 26-SEP-2000; 2000US-0235451P.  
 XX  
 PA (GETH ) GENENTECH INC.  
 XX  
 PI Devaux B, Eberhard D, Goddard A, Godowski PJ, Grimaldi JC;

PI Hillan KJ, Matarabe CK, Wood WI, Yansura DG, Zhang Z;  
 XX WPI, 2002-383270/41.  
 DR N-PSDB; ABR0391.  
 XX  
 PT New polypeptide termed protein upregulated in metastatic prostate cancer  
 PT and encoding polynucleotides, useful for identifying polypeptide  
 PT antigens for treating prostate cancer.

XX Claim 23, Fig 2, 137pp; English.

XX The invention relates to an isolated human protein upregulated in  
 CC metastatic prostate cancer (PMPCn). PRO23203 polypeptide, a sequence 80%  
 CC identical to PRO23203 and the sequence as encoded by cDNA insert of the  
 CC vector deposited as ATCC Deposit No. PTA-2513 (DNA185171-2994) on  
 CC 26/9/2000. Also included are the polynucleotide encoding the protein (or  
 CC a DNA sequence 80% identical to the polynucleotide and one that  
 CC hybridises to complement of the polynucleotide), a vector comprising the  
 CC polynucleotide, a polynucleotide deposited with ATCC under accession  
 CC number PTA-2513 (DNA185171-2994), a host cell comprising the vector,  
 CC preparation of PRO23203, a chimeric molecule comprising PRO23203 fused to  
 CC a heterologous amino acid sequence, an anti-PRO23203 antibody, an  
 CC agonist/antagonist of PRO23203, and diagnosing the presence of prostate  
 CC cancer in a mammal by: (a) contacting a microarray diagnostic with a  
 CC DNA185171-2994 probe, detecting and quantifying hybridisation of  
 CC tissue and determining if DNA185171-2994 is overexpressed; or (b)  
 CC detecting a tissue of the mammal with an anti-PRO23203 antibody and  
 CC detecting the binding of the antibody to a component of the tissue, where  
 CC binding is indicative of the presence of prostate cancer in the mammal.  
 CC The antibody is useful for treating prostate cancer in mammal which is  
 CC androgen independent prostate cancer, that has metastasised to another  
 CC portion of the body, where the antibody is not conjugated with a  
 CC cytotoxic agent and the method further comprises administering a  
 CC chemotherapeutic agent to the mammal and for diagnosing the presence of  
 CC prostate cancer in a mammal. The PRO23203 polynucleotide is useful for  
 CC generating either transgenic animals or knock out animals which in turn  
 CC are useful in the development and screening of therapeutically useful  
 CC reagents or for use in gene therapy and for chromosome identification.  
 CC PRO23203 and polynucleotide may also be used for tissue typing and  
 CC PRO23203 may also be employed as a therapeutic agent and for screening  
 CC compounds to identify those that mimic the PRO23203 polypeptides  
 CC (agonists) or prevent the effect of the PRO23203 polypeptide  
 CC (antagonists). PRO23203 is further useful for the affinity purification  
 CC of PRO23203 from recombinant cell culture or natural sources. The present  
 CC sequence represents PRO23203

XX Sequence 454 AA;

Query Match 100.0%; Score 2351; DB 5; Length 454;  
 Best Local Similarity 100.0%; Pred. No. 5e-245;  
 Matches 454; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MESISMGSPKSLSECTCPNGINGIKDARKVAVGIGSDPFAKSLTIRLIRGCVVVG 60  
 DB 1 MESISMGSPKSLSECTCPNGINGIKDARKVAVGIGSDPFAKSLTIRLIRGCVVVG 60  
 QY 61 RNPKTASEFPFHVVDVTHEDALTNTNIIFVAIHREHYTSLMDLRLHVGKILDIVSNM 120  
 DB 61 RNPKTASEFPFHVVDVTHEDALTNTNIIFVAIHREHYTSLMDLRLHVGKILDIVSNM 120  
 QY 121 RINQVPSNAEYLAIFPDLIVKGNVVSAAWALQGPDAAROVYICSNNOARQCYIE 180  
 DB 121 RINQVPSNAEYLAIFPDLIVKGNVVSAAWALQGPDAAROVYICSNNOARQCYIE 180  
 QY 181 IARQINFIPIIDIGSLSSAREINLPLRFTLRGVPVVAISLATFFLYSFVRDVIHPYA 240  
 DB 181 IARQINFIPIIDIGSLSSAREINLPLRFTLRGVPVVAISLATFFLYSFVRDVIHPYA 240  
 QY 241 RNOQSPFYKIPLEIVNKLPIVAITLISVYLAAGLAAAYOYVGTGTRRRPPEMETLQ 300  
 DB 241 RNOQSPFYKIPLEIVNKLPIVAITLISVYLAAGLAAAYOYVGTGTRRRPPEMETLQ 300

QY 301 CRKQGLISFPFPMVAVVSLCLPMRSEKYLFLNNAVQOVHANINSEWEEVRIEM 360  
 DB 301 CRKQGLISFPFPMVAVVSLCLPMRSEKYLFLNNAVQOVHANINSEWEEVRIEM 360  
 QY 361 ISFGIMSLGLSLAATSVISVSNALNWRPFSTQSTLGVALLISTPHVLIYGMRAPE 420  
 DB 361 ISFGIMSLGLSLAATSVISVSNALNWRPFSTQSTLGVALLISTPHVLIYGMRAPE 420  
 QY 421 EBYRFTYPPNFVALVLPISIVILDLQLCRYPD 454  
 DB 421 EBYRFTYPPNFVALVLPISIVILDLQLCRYPD 454

# RESULT 6

ID AB08893 standard; protein; 454 AA.

AC AB08893;

DT 09-OCT-2003 (first entry)

DE Tumour-associated antigenic target 138 (TAT138).

XX Human; tumour-associated antigenic target polypeptide; TAT; tumour;

XX gene therapy; cytostatic; tissue typing; prostate tumour; cancer.

OS Homo sapiens.

XX US2003060612-A1.

PD 27-MAR-2003.

XX 22-JUN-2001; 2001US-00888257.

PR 28-OCT-1997; 97US-0063540P.

PR 17-JUN-1998; 98US-0089653P.

PR 10-SEP-1998; 98US-0097972P.

PR 08-OCT-1998; 98US-0103678P.

PR 02-JUN-1999; 99US-01032252.

PR 01-SEP-1999; 99US-01032252.

PR 18-FEB-2000; 2000US-0004342.

PR 02-MAR-2000; 2000US-0005841.

PR 30-MAR-2000; 2000US-0008439.

PR 24-AUG-2000; 2000US-02023328.

PR 26-SEP-2000; 2000US-0235451P.

PR 01-DEC-2000; 2000US-02362678.

PR 28-FEB-2001; 2001US-0006520.

PR 01-MAR-2001; 2001US-0006666.

XX (GENTH ) GENENTECH INC.

XX Goddard A, Godowski PJ, Gurney AL, Hillan KJ, Polakis P, Smith V;

PI Wood WI, Wu TD, Zhang Z;

XX WPI; 2003-576355/54.

DR N-PSDB; ACD25893.

XX New nucleic acid, useful for the manufacture of a medicament for

PT diagnosing or treating tumor in a mammal.

PS Claim 12; Fig 10; 71pp; English.

Not prior art

XX The invention discloses human nucleic acids encoding tumour-associated  
 CC antigenic target (TAT) polypeptides, with or without their associated  
 CC signal peptide. Also disclosed is an antibody that specifically binds to  
 CC the TAT polypeptides, a method for detecting the presence of a tumour in  
 CC a mammal and a method for killing a cancer cell expressing the TAT  
 CC polypeptides. The nucleotide sequences are useful in preparing TAT  
 CC polypeptides by recombinant techniques and in gene therapy (e.g. for  
 CC replacement of defective gene). The TAT polypeptides are useful as  
 CC therapeutic agents and for detecting the presence, prevention and/or  
 CC treatment of a tumour, such as colon, breast or prostate tumour. The TAT  
 CC polypeptides and nucleic acids may also be used diagnostically for tissue

CC Typing. The sequence presented is the TARI38 polypeptide of the invention  
 XX Sequence 454 AA;

Query Match 100.0%; Score 2351; DB 7; Length 454;  
 Best Local Similarity 100.0%; Pred. No.5e-245;  
 Matches 454; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MESISMWSPKSLSETCLPENGINGIDARKVTYGVIGSGDFAKSLTIRLRGCHVYVIGS 60  
 DB 1 MESISMWSPKSLSETCLPENGINGIDARKVTYGVIGSGDFAKSLTIRLRGCHVYVIGS 60  
 QY 61 RNKFASEFPFHVDVTHHEDALTKNIIIVAIHREHYTSLMDRLHVGKILIDVSNM 120  
 DB 61 RNKFASEFPFHVDVTHHEDALTKNIIIVAIHREHYTSLMDRLHVGKILIDVSNM 120  
 QY 121 RINQYPSNAEYIASLPDLSLYKGFNVSAVALQGPDAASQVYICSNNIQARQVIE 180  
 DB 121 RINQYPSNAEYIASLPDLSLYKGFNVSAVALQGPDAASQVYICSNNIQARQVIE 180  
 QY 181 LARQNLFIPIIDGLSLSAREIENLPLRFTLMRGPVVAISLATFFLYSPRDVIHPYA 240  
 DB 181 LARQNLFIPIIDGLSLSAREIENLPLRFTLMRGPVVAISLATFFLYSPRDVIHPYA 240  
 QY 241 RNOQSDPYKIPRIYVNTLPIVAITLLSLVYLAGLAAAYQVYGYKRRFPPLMTWQ 300  
 DB 241 RNOQSDPYKIPRIYVNTLPIVAITLLSLVYLAGLAAAYQVYGYKRRFPPLMTWQ 300  
 QY 301 CRKQGLISFPFAMVAVASLCIPMRSSRYFLNMAVQOVHANIENSNNEEVRIRIEM 360  
 DB 301 CRKQGLISFPFAMVAVASLCIPMRSSRYFLNMAVQOVHANIENSNNEEVRIRIEM 360  
 QY 361 ISFGIMSLGLSLAATVSPVSANLNMREFSFIOSTLGYVALLISTFHVLLYGMKRAFE 420  
 DB 361 ISFGIMSLGLSLAATVSPVSANLNMREFSFIOSTLGYVALLISTFHVLLYGMKRAFE 420  
 QY 421 EBYRYFTPPNFVALVPSIVILLDLQLCRYPD 454  
 DB 421 EBYRYFTPPNFVALVPSIVILLDLQLCRYPD 454

RESULT 7  
 ADB65001  
 ID ADB65001 standard; protein; 454 AA.

XX ADB65001;  
 XX 04-DEC-2003 (first entry)

DE Human protein encoded by clone PROST20168600.

XX Human; pharmaceutical; diagnostic; gene therapy; tissue regeneration;  
 KM cell regeneration; membrane protein; signal transduction-related protein;  
 KM transcription-related protein; osteoporosis; neurological disease;  
 KM cancer; tumour.

XX Homo sapiens.

XX EP1308459-A2.

XX 07-MAY-2003.

XX 28-MAR-2002; 2002EP-00007401.

XX 05-NOV-2001; 2001JP-00379298.

XX 25-JAN-2002; 2002US-00350978.

XX (HELI-) HELIX RES INST.

XX (REAS-) RES ASSOC BIOTECHNOLOGY.

PI Iisogai T, Sugiyama T, Otsuki T, Wakamatsu A, Sato H, Ishii S,  
 PI Yamamoto J, Isono Y, Hio Y, Otsuka K, Nagai K, Tamechika I,  
 PI Seki N, Yoshikawa T, Otsuka M, Nagahari K, Masuno Y;

XX WPI; 2003-450961/43.  
 DR N-PSDB; ADB63031.

PT New polynucleotides and polypeptides, useful for developing a diagnostic  
 marker or medicines for regulation of their expression and activity, or  
 as targets of gene therapy.

XX Claim 1; Page; 22pp; English.

The invention discloses a polynucleotide comprising a sequence selected from 1970 fully defined nucleotide sequences which encode novel polypeptides. Also claimed is a polypeptide encoded by the polynucleotide or its partial peptide, an antibody binding to the polypeptide or peptide of the polynucleotide, immunologically assaying the polypeptide or peptide of the polynucleotide by contacting the polypeptide or peptide with the antibody of the encoded protein, and observing the binding between the two, a transformant carrying the polynucleotide in an expressible manner and an antisense polynucleotide. The oligonucleotide is useful as a primer for synthesizing the polynucleotide, or as a probe for detecting the polynucleotide. The polynucleotide and encoded proteins are useful as pharmaceutical agents and many disease-related genes may be included in them, for developing a diagnostic marker or medicines for regulation of their expression and activity, or as targets of gene therapy. The genes are involved in tissue and/or cell regeneration. Membrane proteins, signal transduction-related proteins, transcription-related proteins, disease-related proteins and genes encoding them can be used as indicators for diseases (e.g. osteoporosis, neurological diseases, cancer, tumours). The cDNA may be used to regulate the activity or expression of the encoded protein to treat diseases. The sequence presented is a protein of the invention. Note: Some of the sequence data for this patent is not represented in the printed specification, but is based on sequence information supplied by the European Patent Office.

XX Sequence 454 AA;

Query Match 99.9%; Score 2348; DB 7; Length 454;  
 Best Local Similarity 99.8%; Pred. No.1.e-244;  
 Matches 453; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MESISMWSPKSLSETCLPENGINGIDARKVTYGVIGSGDFAKSLTIRLRGCHVYVIGS 60  
 DB 1 MESISMWSPKSLSETCLPENGINGIDARKVTYGVIGSGDFAKSLTIRLRGCHVYVIGS 60  
 QY 61 RNKFASEFPFHVDVTHHEDALTKNIIIVAIHREHYTSLMDRLHVGKILIDVSNM 120  
 DB 61 RNKFASEFPFHVDVTHHEDALTKNIIIVAIHREHYTSLMDRLHVGKILIDVSNM 120  
 QY 121 RINQYPSNAEYIASLPDLSLYKGFNVSAVALQGPDAASQVYICSNNIQARQVIE 180  
 DB 121 RINQYPSNAEYIASLPDLSLYKGFNVSAVALQGPDAASQVYICSNNIQARQVIE 180  
 QY 181 LARQNLFIPIIDGLSLSAREIENLPLRFTLMRGPVVAISLATFFLYSPRDVIHPYA 240  
 DB 181 LARQNLFIPIIDGLSLSAREIENLPLRFTLMRGPVVAISLATFFLYSPRDVIHPYA 240  
 QY 241 RNOQSDPYKIPRIYVNTLPIVAITLLSLVYLAGLAAAYQVYGYKRRFPPLMTWQ 300  
 DB 241 RNOQSDPYKIPRIYVNTLPIVAITLLSLVYLAGLAAAYQVYGYKRRFPPLMTWQ 300  
 QY 301 CRKQGLISFPFAMVAVASLCIPMRSSRYFLNMAVQOVHANIENSNNEEVRIRIEM 360  
 DB 301 CRKQGLISFPFAMVAVASLCIPMRSSRYFLNMAVQOVHANIENSNNEEVRIRIEM 360  
 QY 361 ISFGIMSLGLSLAATVSPVSANLNMREFSFIOSTLGYVALLISTFHVLLYGMKRAFE 420  
 DB 361 ISFGIMSLGLSLAATVSPVSANLNMREFSFIOSTLGYVALLISTFHVLLYGMKRAFE 420  
 QY 421 EBYRYFTPPNFVALVPSIVILLDLQLCRYPD 454  
 DB 421 EBYRYFTPPNFVALVPSIVILLDLQLCRYPD 454

RESULT 8  
AAE28951 standard; protein; 490 AA.  
ID AAE28951 standard; protein; 490 AA.  
XX  
AC AAE28951;  
XX  
DT 27-JAN-2003 (first entry)  
XX  
DE Human STEAPRP protein.  
XX  
KM Human; STEAPRP; proliferative disorder; prostate hyperplasia; cancer;  
KM six-transmembrane epithelial antigen of the prostate related protein;  
KM gene therapy; cytostatic.  
XX  
OS Homo sapiens.  
XX  
PN WO200272596-A1.  
XX  
PD 19-SEP-2002.  
XX  
PF 07-MAR-2002; 2002WO-US007053.  
XX  
PR 09-MAR-2001; 2001US-00802520.  
XX  
PA (INCY-) INCYTE GENOMICS INC.  
XX  
PI Lal PG, Faris M, Chen H, Ison CH;  
XX  
DR WPI; 2002-723311/78.  
DR N-PSDB; AAD46429.  
XX  
PT New six-transmembrane epithelial antigen of the prostate-related protein  
PT and nucleic acids, useful for diagnosing and treating prostate cell  
PT proliferative disorders, particularly prostate hyperplasia and prostate  
PT cancer.  
XX  
PS Claim 13; Page 54-55; 62zp; English.  
XX  
CC The invention relates to six-transmembrane epithelial antigen of the  
CC prostate related protein (STEAPRP) and its corresponding nucleic acid.  
CC STEAPRP DNA is used to diagnose and to treat prostate cell proliferative  
CC disorders, particularly prostate hyperplasia and prostate cancer, and for  
CC screening a library of molecules of compounds for specific binding  
CC affinity. It is also used in gene therapy. STEAPRP is used to screen a  
CC subject sample for antibodies, which specifically binds the protein, and  
CC to prepare and purify a protein. The present sequence is human STEAPRP  
CC protein  
XX  
SQ Sequence 490 AA;

Query Match 97.6%; Score 2294; DB 5; Length 490;  
Best Local Similarity 100.0%; Pred. No. 8.4e-239;  
Matches 444; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MESISWMSGSPKSLSTCTCPNGINGIKKARKTGVYIGSGDPKSLTIRLRGCHVYIGS 60  
DB 1 MESISWMSGSPKSLSTCTCPNGINGIKKARKTGVYIGSGDPKSLTIRLRGCHVYIGS 60  
QY 61 RNPKEASEFPFHVVDTVTHEDALTKTNIIFVAIHEHRTSLAMDRLHVLGKILDIVNNM 120  
DB 61 RNPKEASEFPFHVVDTVTHEDALTKTNIIFVAIHEHRTSLAMDRLHVLGKILDIVNNM 120  
QY 121 RINQYPSNAEYLASLPDLSLIVGFNVVSAMALQGLPKDSRQVYICSNNIQARQVIE 180  
DB 121 RINQYPSNAEYLASLPDLSLIVGFNVVSAMALQGLPKDSRQVYICSNNIQARQVIE 180  
QY 181 LABQINFPIDLGSISSAREIENPLRLFTLMRGVVAISLAFEFVSRVVRVHPEVA 240  
DB 181 LABQINFPIDLGSISSAREIENPLRLFTLMRGVVAISLAFEFVSRVVRVHPEVA 240  
QY 241 RNOQSDPYKIPLEIVNKTLPVIAITLLSLVYIAGLAAAYQLYGKRRPFWLTWQ 300

Db 241 RNOQSDPYKIPLEIVNKTLPVIAITLLSLVYIAGLAAAYQLYGKRRPFWLTWQ 300  
QY 301 CRKQGLISFPFAMVAVAYSLCPMRRSERYLFLNNAVOOVANIENSNNEEVRMIEM 360  
DB 301 CRKQGLISFPFAMVAVAYSLCPMRRSERYLFLNNAVOOVANIENSNNEEVRMIEM 360  
QY 361 ISFGIMSLGLSLANTSTPSVSNALNMBEBSFIOSTLGVALLSTFHYLYGKRAE 420  
DB 361 ISFGIMSLGLSLANTSTPSVSNALNMBEBSFIOSTLGVALLSTFHYLYGKRAE 420  
QY 421 EBYRFRYTPPNFVALVLPISIVL 444  
DB 421 EBYRFRYTPPNFVALVLPISIVL 444  
RESULT 9  
ABU63312 standard; protein; 490 AA.  
ID ABU63312 standard; protein; 490 AA.  
XX  
AC ABU63312;  
XX  
DT 18-SEP-2003 (first entry)  
XX  
DE Human transmembrane protein TMPL.  
XX  
KM Human; prostate tumour; lung tumour; TMPL; cancer.  
XX  
OS Homo sapiens.  
XX  
PN US2003064397-A1.  
XX  
PD 03-APR-2003.  
XX  
PF 24-JUL-2002; 2002US-00205267.  
XX  
PR 22-MAY-1998; 98US-00083521.  
PR 16-SEP-1999; 99US-00397558.  
PR 09-MAR-2001; 2001US-00802520.  
PR 26-SEP-2001; 2001US-00963896.  
XX  
PA (INCY-) INCYTE GENOMICS INC.  
XX  
PI Spancake KM, Rickert PK, Lal PG, Ison CH;  
XX  
DR WPI; 2003-540783/51.  
DR N-PSDB; ACD27642.  
XX  
PT New transmembrane protein differentially expressed in prostate and lung  
PT tumors, useful for diagnosing, staging, treating or monitoring  
PT progression or treatment of cancer, preferably lung or prostate cancer.  
XX  
PS Claim 1; Fig 1; 49zp; English.  
XX  
CC The invention relates to a purified human protein comprising a  
CC polypeptide having a fully defined transmembrane protein differentially  
CC expressed in prostate and lung tumour (TMPL) sequence. The protein is  
CC useful for detecting expression of the protein in a sample, where the  
CC sample is lung or prostate tissue sample and the protein is  
CC differentially expressed when compared with a standard and is diagnostic  
CC of a lung or prostate cancer. The protein is useful to identify an  
CC antibody that specifically binds the protein. The protein is useful for  
CC screening several molecules and compounds to identify at least one  
CC ligand. The protein is also useful for preparing and purifying a  
CC polyclonal antibody and for preparing a monoclonal antibody using  
CC standard animal immunisation/hybridoma methods. A TMPL specific antibody  
CC is useful to detect expression of the protein in a lung or prostate  
CC tissue sample and the method is useful as a diagnostic for lung or  
CC prostate cancer. A TMPL specific antibody is also useful for  
CC immunoprecipitating a protein, for treating a prostate cancer and for  
CC delivering a therapeutic agent to a cancer, preferably prostate cancer.  
CC The protein is useful for diagnosing, staging, treating or monitoring  
CC progression or treatment of cancer, preferably lung or prostate cancer.  
CC The present sequence represents the amino acid sequence of the human

CC transmembrane protein differentially expressed in prostate and lung  
 CC tumour TMPL  
 XX  
 SQ Sequence 490 AA;

Query Match 97.6%; Score 2294; DB 7; Length 490;  
 Best Local Similarity 100.0%; Pred. No. 8,4e-239; Indels 0; Gaps 0;  
 Matches 443; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MESISWMSGSPKSISETCLPENGINGIDKARKVTGVIGSGDPKASLTIRLRGCHVYIGS 60  
 DB 1 MESISWMSGSPKSISETCLPENGINGIDKARKVTGVIGSGDPKASLTIRLRGCHVYIGS 60  
 QY 61 RNKFASEFPFHVVDTTHEDALTNTNIIIVAIHREHYTSLMDLRHLVSKILLIDVSNM 120  
 DB 61 RNKFASEFPFHVVDTTHEDALTNTNIIIVAIHREHYTSLMDLRHLVSKILLIDVSNM 120  
 QY 121 RINQPESSNAEYIASLPDLSLYKGFNVVSAMALQGPDKASROYICSNNIQARQOYIE 180  
 DB 121 RINQPESSNAEYIASLPDLSLYKGFNVVSAMALQGPDKASROYICSNNIQARQOYIE 180  
 QY 181 LARQNLPIPDLSGSARIEINLPRLFTLMRGPVVAISLATPFLYSFVDVTHPYA 240  
 DB 181 LARQNLPIPDLSGSARIEINLPRLFTLMRGPVVAISLATPFLYSFVDVTHPYA 240  
 QY 241 RNOQSDPYKPIPIEIVNKTPIVAITLISLYAGLAAAYQLYGTYKRRFPFWLETLQ 300  
 DB 241 RNOQSDPYKPIPIEIVNKTPIVAITLISLYAGLAAAYQLYGTYKRRFPFWLETLQ 300  
 QY 301 CRKQGLISFPFPAVAVHAYSLCLPMRSERYFLNMAVQOVANINENSNNEEVRIRIEM 360  
 DB 301 CRKQGLISFPFPAVAVHAYSLCLPMRSERYFLNMAVQOVANINENSNNEEVRIRIEM 360  
 QY 361 ISFGIMSLGLSLAATSIIPSVSNALNMRFSFIQSTLGYVALLISTFVLLIYGWKRAFE 420  
 DB 361 ISFGIMSLGLSLAATSIIPSVSNALNMRFSFIQSTLGYVALLISTFVLLIYGWKRAFE 420  
 QY 421 EBYRYFTPPNFVTLALVLPISIVL 444  
 DB 421 EBYRYFTPPNFVTLALVLPISIVL 444

RESULT 10  
 AAU0187  
 ID AAU0187 standard; protein; 490 AA.  
 AC AAU0187;

DT 16-JAN-2002 (first entry)

DE Human Six-Transmembrane Protein of Prostate 1, STMPL.

XX Human; Six-Transmembrane Protein of Prostate 1; STM, 1; prostate cancer;  
 KM benign prostatic hyperplasia; acute prostatic; testicular cancer;  
 KM cryptorchidism; testicular disorder; proliferative disorder; lymphoma;  
 KM leukemia; melanoma; ovarian cancer; breast cancer; pancreatic cancer;  
 KM liver cancer; lung cancer; cytostatic.

XX Homo sapiens.

XX Key location/Qualifiers

FT 209..230 /label= Transmembrane\_domain\_1

FT 255..273 /label= Transmembrane\_domain\_2

FT 304..325 /label= Transmembrane\_domain\_3

FT 360..380 /label= Transmembrane\_domain\_4

FT 393..415 /label= Transmembrane\_domain\_5

FT 432..452 /label= Transmembrane\_domain\_6

XX W0200172962-A2.  
 FM 04-OCT-2001.  
 XX 23-MAR-2001; 2001WO-US009410.  
 XX 24-MAR-2000; 2000US-0191929P.

XX (SAAT/) SAATCLOGLU F.

XX Saatcloglu F;

XX WPI: 2001-662926/76.

XX N-PSDE: AAS15793, AAS15801, AAS15802.

PT New polynucleotide for the diagnosis, prevention and treatment for  
 prostate and testis disorders, particularly prostate cancer, comprises  
 prostate-specific or testis-specific nucleic acids.

PS Claim 1; Fig 4B; 114pp; English.

CC The invention relates to substantially pure prostate-specific or testis-  
 specific polypeptides and the nucleic acids encoding them. Also included  
 are vectors and host cells expressing the proteins, a transgenic animal  
 expressing the protein, antibodies against the proteins, probes for  
 detecting the nucleic acids, antisense molecules for the nucleic acids  
 and methods of isolating modulators of the proteins. Compounds that  
 modulate the prostate specific or testis specific polypeptide are useful  
 to diagnose, prevent or treat disorders of the testis or prostate  
 particularly prostate cancer, benign prostatic hyperplasia, acute  
 prostatitis, testicular cancer, cryptorchidism, undescended, retractile,  
 ascending or vanished testis. Other proliferative disorders for which the  
 modulators may be used include lymphoma, leukaemia, melanoma, ovarian  
 cancer, breast cancer, pancreatic cancer, liver cancer and lung cancer.  
 CC The present sequence represents a prostate specific protein, Six-  
 transmembrane Protein of Prostate 1, STM,1

CC Sequence 490 AA;

Query Match 97.4%; Score 2290; DB 4; Length 490;  
 Best Local Similarity 99.8%; Pred. No. 2.3e-238;  
 Matches 443; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MESISWMSGSPKSISETCLPENGINGIDKARKVTGVIGSGDPKASLTIRLRGCHVYIGS 60  
 DB 1 MESISWMSGSPKSISETCLPENGINGIDKARKVTGVIGSGDPKASLTIRLRGCHVYIGS 60  
 QY 61 RNKFASEFPFHVVDTTHEDALTNTNIIIVAIHREHYTSLMDLRHLVSKILLIDVSNM 120  
 DB 61 RNKFASEFPFHVVDTTHEDALTNTNIIIVAIHREHYTSLMDLRHLVSKILLIDVSNM 120  
 QY 121 RINQPESSNAEYIASLPDLSLYKGFNVVSAMALQGPDKASROYICSNNIQARQOYIE 180  
 DB 121 RINQPESSNAEYIASLPDLSLYKGFNVVSAMALQGPDKASROYICSNNIQARQOYIE 180  
 QY 181 LARQNLPIPDLSGSARIEINLPRLFTLMRGPVVAISLATPFLYSFVDVTHPYA 240  
 DB 181 LARQNLPIPDLSGSARIEINLPRLFTLMRGPVVAISLATPFLYSFVDVTHPYA 240  
 QY 241 RNOQSDPYKPIPIEIVNKTPIVAITLISLYAGLAAAYQLYGTYKRRFPFWLETLQ 300  
 DB 241 RNOQSDPYKPIPIEIVNKTPIVAITLISLYAGLAAAYQLYGTYKRRFPFWLETLQ 300  
 QY 301 CRKQGLISFPFPAVAVHAYSLCLPMRSERYFLNMAVQOVANINENSNNEEVRIRIEM 360  
 DB 301 CRKQGLISFPFPAVAVHAYSLCLPMRSERYFLNMAVQOVANINENSNNEEVRIRIEM 360  
 QY 361 ISFGIMSLGLSLAATSIIPSVSNALNMRFSFIQSTLGYVALLISTFVLLIYGWKRAFE 420  
 DB 361 ISFGIMSLGLSLAATSIIPSVSNALNMRFSFIQSTLGYVALLISTFVLLIYGWKRAFE 420  
 QY 421 EBYRYFTPPNFVTLALVLPISIVL 444



DB 421 EHYRFTPPNVLALVPSIVL 444

RESULT 11  
AAU10189  
ID AAU10189 standard; protein; 419 AA.

AC AAU10189;  
XX  
XX  
XX 16-JAN-2002 (first entry)  
XX  
XX  
XX Human ORF3 of Six-Transmembrane Protein of Prostate 1.  
DE  
XX  
XX Human; Six-Transmembrane Protein of Prostate 1, STM, 1; prostate cancer;  
XX benign prostatic hyperplasia; acute prostatitis; testicular cancer;  
XX cryptorchidism; testicular disorder; proliferative disorder; lymphoma;  
XX leukemia; melanoma; ovarian cancer; breast cancer; pancreatic cancer;  
XX liver cancer; lung cancer; cystostatic; ORF3.  
XX  
XX Homo sapiens.  
OS  
XX  
XX WO200172962-A2.  
XX  
XX 04-OCT-2001.  
XX  
XX 23-MAR-2001; 2001WO-US009410.  
XX  
XX 24-MAR-2000; 2000US-0191929P.  
XX  
XX (SAAT/) SAATCLOGLU F.  
XX  
XX Saactcloglu F;  
XX  
XX WPI: 2001-662926/76.  
XX  
XX N-PSDB; AAS15813.

PT New polynucleotide for the diagnosis, prevention and treatment for  
PT prostate and testis disorders, particularly prostate cancer, comprises  
PT prostate-specific or testis-specific nucleic acids.  
XX  
XX Claim 1; Fig 4K; 114pp; English.

XX  
XX The invention relates to substantially pure prostate-specific or testis-  
XX specific polypeptides and the nucleic acids encoding them. Also included  
XX are vectors and host cells expressing the proteins, a transgenic animal  
XX expressing the protein, antibodies against the proteins, probes for  
XX detecting the nucleic acids, antisense molecules for the nucleic acids  
XX and methods of isolating modulators of the proteins. Compounds that  
XX modulate the prostate specific or testis specific polypeptide are useful  
XX to diagnose, prevent or treat disorders of the testis or prostate  
XX particularly prostate cancer, benign prostatic hyperplasia, acute  
XX prostatitis, testicular cancer, cryptorchidism, undescended, retractile,  
XX ascending or vanished testis. Other proliferative disorders for which the  
XX modulators may be used include lymphoma, leukemia, melanoma, ovarian  
XX cancer, breast cancer, pancreatic cancer, liver cancer and lung cancer.  
XX  
XX The present sequence is prostate specific protein, Six-Transmembrane  
XX Protein of Prostate 1, STM,1, ORF3  
XX  
XX  
XX Sequence 419 AA;

Query Match 86.6%; Score 2036; DB 4; Length 419;  
Best Local Similarity 99.7%; Pred. No. 5.5e-211;  
Matches 394; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MESISMSPKSLSETCPLNGINGIKDARKTVGVIGSGDPKSLTRLIRCGHVVIGS 60  
DB 1 MESISMSPKSLSETCPLNGINGIKDARKTVGVIGSGDPKSLTRLIRCGHVVIGS 60  
QY 61 RNPFRASEFPFHVVDVTHEDALTKNTIIFVAIRREHYTSIMDLRHLLVGLKILIVSNM 120  
DB 61 RNPFRASEFPFHVVDVTHEDALTKNTIIFVAIRREHYTSIMDLRHLLVGLKILIVSNM 120

QY 121 RINQYESNAEYIASLFPDLSLVKGFNVVSAMALQIGPDASROYTICSNNTIARQVTE 180  
DB 121 RINQYESNAEYIASLFPDLSLVKGFNVVSAMALQIGPDASROYTICSNNTIARQVTE 180  
QY 181 IARQLNFIPIIDGSSAREINLPLRLFTLRGPPVVAISLATPELXSPRDVHPYA 240  
DB 181 IARQLNFIPIIDGSSAREINLPLRLFTLRGPPVVAISLATPELXSPRDVHPYA 240  
QY 241 RNOQSDPYKIPLEIVNKTLPVIAITLLSLVYLAGLLAAAYQDYGTXYRPPPWLETWLQ 300  
DB 241 RNOQSDPYKIPLEIVNKTLPVIAITLLSLVYLAGLLAAAYQDYGTXYRPPPWLETWLQ 300  
QY 301 CRKQGLSLFPFAMVVAASLCLPMRBSRYLFLMAAYOYVANTENSWNEEVRIMY 360  
DB 301 CRKQGLSLFPFAMVVAASLCLPMRBSRYLFLMAAYOYVANTENSWNEEVRIMY 360  
QY 361 ISFGIMSLGILSLAVTSIPSVNMLMWEPFIIQ 395  
DB 361 ISFGIMSLGILSLAVTSIPSVNMLMWEPFIIQ 395

RESULT 12  
ABG12306  
ID ABG12306 standard; protein; 576 AA.

AC ABG12306;  
XX  
XX 18-FEB-2002 (first entry)  
XX  
XX Novel human diagnostic protein #12297.  
XX  
XX Human; chromosome mapping; gene mapping; gene therapy; forensic;  
XX food supplement; medical imaging; diagnostic; genetic disorder.  
XX  
XX Homo sapiens.  
OS  
XX  
XX WO200175067-A2.  
XX  
XX 11-OCT-2001.  
XX  
XX 30-MAR-2001; 2001WO-US008631.  
XX  
XX 31-MAR-2000; 2000US-00540217.  
XX  
XX 23-AUG-2000; 2000US-00649167.  
XX  
XX (HYSE-) HYSEQ INC.  
XX  
XX Dmanac RT, Liu C, Tang YT;  
XX  
XX WPI: 2001-639362/73.  
XX  
XX N-PSDB; AAS76493.

PT New isolated polynucleotide and encoded polypeptides, useful in  
PT diagnostics, forensics, gene mapping, identification of mutations  
PT responsible for genetic disorders or other traits and to assess  
PT biodiversity.  
XX  
XX Claim 20; SEQ ID NO 42665; 103pp; English.

XX  
XX The invention relates to isolated polynucleotide (I) and polypeptide (II)  
XX sequences. (I) is useful as hybridisation probes, polymerase chain  
XX reaction (PCR) primers, oligomers, and for chromosome and gene mapping,  
XX and in recombinant production of (II). The polynucleotides are also used  
XX in diagnostics as expressed sequence tags for identifying expressed  
XX genes. (I) is useful in gene therapy techniques to restore normal  
XX activity of (II) or to treat disease states involving (II). (II) is  
XX useful for generating antibodies against it, detecting or quantitating a  
XX polypeptide in tissue, as molecular weight markers and as a food  
XX supplement. (II) and its binding partners are useful in medical imaging  
XX of sites expressing (II). (I) and (II) are useful for treating disorders  
XX involving aberrant protein expression or biological activity. The  
XX polypeptide and polynucleotide sequences have applications in  
XX diagnostics, forensics, gene mapping, identification of mutations

CC responsible for genetic disorders or other traits to assess biodiversity  
CC and to produce other types of data and products dependent on DNA and  
CC amino acid sequences. ABG00010-ABG0377 represent novel human diagnostic  
CC patent did not appear in the printed specification, but was obtained in  
CC electronic format directly from WIPO at  
CC ftp.wipo.int/pub/published\_pct\_sequences  
XX

SO Sequence 576 AA;

Query Match 83.9%; Score 1972.5; DB 4; Length 576;  
Best Local Similarity 87.4%; Pred. No. 6.7e-204;  
Matches 390; Conservative 0; Mismatches 1; Indels 55; Gaps 1;

QY 1 MESISMWSPKSLSEFCPLNGINGIKDARKTVGVIGSDPKSLTIRIRCGYHVIGS 60  
DB 1 MESISMWSPKSLSEFCPLNGINGIKDARKTVGVIGSDPKSLTIRIRCGYHVIGS 60  
QY 61 RNPKFASSEFPFHVDVTHHEDALTNTNII FVAIHREHYTSLMDLRHLVKGKILIDVSNM 120  
DB 61 RNPKFASSEFPFHVDVTHHEDALTNTNII FVAIHREHYTSLMDLRHLVKGKILIDVSNM 120  
QY 121 RINQYPSNAEYLAIFPDSLIYKGFNVSAVALQGPDSARQVYICSNNTQARQVIE 180  
DB 121 RINQYPSNAEYLAIFPDSLIYKGFNVSAVALQGPDSARQVYICSNNTQARQVIE 180  
QY 181 LARQINFPIDGSLSSAREIENTPLRLFTLRGPPVVAISLATEFFLYSPRVDIHPYA 240  
DB 181 LARQINFPIDGSLSSAREIENTPLRLFTLRGPPVVAISLATEFFLYSPRVDIHPYA 240  
QY 241 RNOQSDPYKPIPIEIVNKTLPYVATLLSLVYLAGLAAAYQLYGKTRRPPMLETWLO 300  
DB 241 RNOQSDPYKPIPIEIVNKTLPYVATLLSLVYLAGLAAAYQLYGKTRRPPMLETWLO 300  
QY 301 CRKQGLSFFPAMVAVASLCLPMRBSERYFLNMAVOQHANIENSMNEEVRRIEMY 360  
DB 301 CRKQGLSFFPAMVAVASLCLPMRBSERYFLNMAVOQHANIENSMNEEVRRIEMY 360  
QY 361 ISFGIMSLGLSLAVTSIPSVSNALNMBEFSFIQSTLGYVALLISTFHVLIYGMKRAFE 420  
DB 361 ISFGIMSLGLSLAVTSIPSVSNALNMBEFSFIQSTLGYVALLISTFHVLIYGMKRAFE 420  
QY 421 EBYRFTYPPNFVALVLPISIVILD 446  
DB 421 EBYRFTYPPNFVALVLPISIVILD 446  
QY 463 EBYRFTYPPNFVALVLPISIVILD 491  
DB 463 EBYRFTYPPNFVALVLPISIVILD 491

RESULT 13

ABG00113  
ID ABG00113 standard; protein; 1273 AA.

AC ABG00113;

DT 13-FEB-2002 (first entry)

DE Novel human diagnostic protein #104.

KM Human; chromosome mapping; gene mapping; gene therapy; forensic;  
KW Food supplement; medical imaging; diagnostic; genetic disorder.

OS Homo sapiens.

PN WO200175067-A2.

PD 11-OCT-2001.

PF 30-MAR-2001; 2001WO-US008631.

PR 31-MAR-2000; 2000US-00540217.

PR 23-AUG-2000; 2000US-00649167.

PA (HISE-) HISEQ INC.

PI Dmanac RT, Liu C, Tang YT;  
XX  
XX WPI; 2001-639362/73.  
DR N-PSDB; AAS64300.  
XX  
PT New isolated polynucleotide and encoded polypeptides, useful in  
PT diagnostics, forensic, gene mapping, identification of mutations  
PT responsible for genetic disorders or other traits and to assess  
PT biodiversity.  
XX  
XX Claim 20; SEQ ID NO 30472; 103pp; English.

CC The invention relates to isolated polynucleotide (I) and polypeptide (II)  
CC sequences. (I) is useful as hybridisation probes, polymerase chain  
CC reaction (PCR) primers, oligomers and for chromosome and gene mapping,  
CC and in recombinant production of (II). The polynucleotides are also used  
CC in diagnostics as expressed sequence tags for identifying expressed  
CC genes. (I) is useful in gene therapy techniques to restore normal  
CC activity of (II) or to treat disease states involving (II). (II) is  
CC useful for generating antibodies against it, detecting or quantitating a  
CC polypeptide in tissue, as molecular weight markers and as a food  
CC supplement. (II) and its binding partners are useful in medical imaging  
CC of sites expressing (II). (I) and (II) are useful for treating disorders  
CC involving aberrant protein expression or biological activity. The  
CC polypeptide and polynucleotide sequences have applications in  
CC diagnostics, forensic, gene mapping, identification of mutations  
CC responsible for genetic disorders or other traits to assess biodiversity  
CC and to produce other types of data and products dependent on DNA and  
CC amino acid sequences. ABG00010-ABG0377 represent novel human diagnostic  
CC amino acid sequences of the invention. Note: The sequence data for this  
CC patent did not appear in the printed specification, but was obtained in  
CC electronic format directly from WIPO at  
CC ftp.wipo.int/pub/published\_pct\_sequences  
XX

SO Sequence 1273 AA;

Query Match 83.9%; Score 1972.5; DB 4; Length 1273;  
Best Local Similarity 87.4%; Pred. No. 2.3e-203;  
Matches 390; Conservative 0; Mismatches 1; Indels 55; Gaps 1;

QY 1 MESISMWSPKSLSEFCPLNGINGIKDARKTVGVIGSDPKSLTIRIRCGYHVIGS 60  
DB 73 MESISMWSPKSLSEFCPLNGINGIKDARKTVGVIGSDPKSLTIRIRCGYHVIGS 132  
QY 61 RNPKFASSEFPFHVDVTHHEDALTNTNII FVAIHREHYTSLMDLRHLVKGKILIDVSNM 120  
DB 133 RNPKFASSEFPFHVDVTHHEDALTNTNII FVAIHREHYTSLMDLRHLVKGKILIDVSNM 192  
QY 121 RINQYPSNAEYLAIFPDSLIYKGFNVSAVALQGPDSARQVYICSNNTQARQVIE 180  
DB 121 RINQYPSNAEYLAIFPDSLIYKGFNVSAVALQGPDSARQVYICSNNTQARQVIE 252  
QY 181 LARQINFPIDGSLSSAREIENTPLRLFTLRGPPVVAISLATEFFLYSPRVDIHPYA 240  
DB 181 LARQINFPIDGSLSSAREIENTPLRLFTLRGPPVVAISLATEFFLYSPRVDIHPYA 312  
QY 241 RNOQSDPYKPIPIEIVNKTLPYVATLLSLVYLAGLAAAYQLYGKTRRPPMLETWLO 300  
DB 241 RNOQSDPYKPIPIEIVNKTLPYVATLLSLVYLAGLAAAYQLYGKTRRPPMLETWLO 372  
QY 301 CRKQGLSFFPAMVAVASLCLPMRBSERYFLNMAVOQHANIENSMNEEVRRIEMY 360  
DB 301 CRKQGLSFFPAMVAVASLCLPMRBSERYFLNMAVOQHANIENSMNEEVRRIEMY 411  
QY 361 ISFGIMSLGLSLAVTSIPSVSNALNMBEFSFIQSTLGYVALLISTFHVLIYGMKRAFE 420  
DB 361 ISFGIMSLGLSLAVTSIPSVSNALNMBEFSFIQSTLGYVALLISTFHVLIYGMKRAFE 437  
QY 421 EBYRFTYPPNFVALVLPISIVILD 446  
DB 421 EBYRFTYPPNFVALVLPISIVILD 446  
QY 463 EBYRFTYPPNFVALVLPISIVILD 491  
DB 463 EBYRFTYPPNFVALVLPISIVILD 491

RESULT 14  
AAE02841  
ID AAE02841 standard; protein; 450 AA.  
XX  
XX AAE02841;  
AC  
XX  
DT 06-AUG-2001 (first entry)  
XX  
XX  
DE Human STEAP-2 protein, alternative version.  
XX  
XX Human; cytosolic; antiproliferative; vaccine; gene therapy;  
KM six transmembrane epithelial antigen of the prostate-2; STEAP-2;  
KW chromosome 7q21; cancer; prostate; colon; bladder; lung; ovarian;  
XX pancreatic.  
XX  
OS Homo sapiens.  
XX  
FH Key Location/Qualifiers  
FT Misc-difference 335..336  
FT /note= "Encoded by GACTGAGCT"  
FT Misc-difference 375..376  
FT /note= "Encoded by AGATGAGT"  
FT Misc-difference 415..416  
FT /note= "Encoded by GAGTAAGC"  
FT Misc-difference 445..446  
FT /note= "Encoded by ACATGAGT"  
FT Misc-difference 448..449  
FT /note= "Encoded by AATTAATTC"  
XX  
XX MO200140276-A2.  
XX  
XX 07-JUN-2001.  
XX  
XX 06-DEC-2000; 2000WO-US033040.  
XX  
XX 06-DEC-1999; 99US-00455486.  
XX  
XX (UROG-) UROGENESYS INC.  
XX  
XX Afar DEH, Hubert RS, Raitano AB, Saffran DC, Mitchell SC,  
PI Faris M, Jakobovits A;  
XX  
XX WPI; 2001-367804/38.  
XX  
XX N-PSDB; ABD07072.  
XX  
XX  
XX New STEAP (six transmembrane epithelial antigen of the prostate)  
PT proteins, expressed in human cancers, useful for detecting and treating  
PT cancer.  
XX  
XX  
XX Claim 1; Page 175-176; 187pp; English.  
XX  
XX The present sequence is an alternative version of human six transmembrane  
CC epithelial antigen of the prostate (STEAP)-2 protein. STEAP-2 is a member  
CC of cell surface serpentine transmembrane antigens. STEAP-2 gene is  
CC located on chromosome 7q21 and is used in gene therapy. Inhibiting the  
CC development or progression of a cancer (eg. prostate, colon, bladder,  
CC lung, ovarian and pancreatic) expressing STEAP or inhibiting growth or  
CC killing cells expressing STEAP in a patient, comprises administering a  
CC vaccine composition to the patient. Treating a patient with a cancer that  
CC expresses STEAP, or inhibiting growth or killing cells expressing STEAP,  
CC comprises administering to the patient a vector encoding single chain  
CC monoclonal antibody that comprises the variable domains of the heavy and  
CC light chains of the monoclonal antibody that specifically binds to STEAP.  
CC such that the vector delivers the single chain monoclonal antibody coding  
CC sequence to the cancer cells and the encoded single chain monoclonal  
CC antibody is expressed intracellularly. Note: This sequence is stated to  
CC be the same as that being shown as SEQ ID NO:8 (AAE02781) in figure 9A-9C  
CC of the specification. However the present sequence lacks several amino  
CC acids at its N-terminal end and has additional amino acids at its C-  
XX terminal end when compared with the sequence shown in figure 9A-9C  
XX  
XX Sequence 450 AA;

Query Match 73.9%; Score 1738; DB 4; Length 450;  
Best Local Similarity 100.0%; Pred No. 1,1e-178; Indels 0; Gaps 0;  
Matches 336; Conservative 0; Mismatches 0;  
QY 119 NMRINQYPPESNAEYLALEPDSLIYGFNVVSAAMALQLGPKASRQVYICSNNIQARQV 178  
DB 1 NMRINQYPPESNAEYLALEPDSLIYGFNVVSAAMALQLGPKASRQVYICSNNIQARQV 60  
QY 179 IEIARQINFPIDIDGLSSARETENLPLFTLMRGPPVVAISLTFEFLYSFVRVIRP 238  
DB 61 IEIARQINFPIDIDGLSSARETENLPLFTLMRGPPVVAISLTFEFLYSFVRVIRP 120  
QY 239 YARNQSDPKYKPIEIVNKTLPVAILTLISLYLAGLAAAYQLYGKYRRFPPLQETW 298  
DB 121 YARNQSDPKYKPIEIVNKTLPVAILTLISLYLAGLAAAYQLYGKYRRFPPLQETW 180  
QY 299 LQCRKQLGLSPFPANVAVYSLCPMRSEERYLPNMAVQOVHANIENSNNEEYWRLE 358  
DB 181 LQCRKQLGLSPFPANVAVYSLCPMRSEERYLPNMAVQOVHANIENSNNEEYWRLE 240  
QY 359 MYISFGIMSLGLSLAATVTSIPSVSNALNMRSEFQSTLGYVALLISTFHYLIYGMKRA 418  
DB 241 MYISFGIMSLGLSLAATVTSIPSVSNALNMRSEFQSTLGYVALLISTFHYLIYGMKRA 300  
QY 419 FEEERYRFPYTPNPFVIALVLPSTIVILDLQLCRYPD 454  
DB 301 FEEERYRFPYTPNPFVIALVLPSTIVILDLQLCRYPD 336  
RESULT 15  
ABP62883  
ID ABP62883 standard; protein; 1082 AA.  
XX  
XX AC ABP62883;  
XX  
XX DT 14-OCT-2002 (first entry)  
XX  
XX  
XX Human polypeptide SEQ ID NO 320.  
XX  
XX  
XX Human; vulnery; dermatological; neuroprotective; nocotropic; cancer;  
KW antiparkinsonian; immunostimulant; cytosolic; immunosuppressive;  
KW antidiabetic; antiallergic; gene therapy; wound healing; tissue repair;  
KW burn; central nervous system disorder; Alzheimer's disease;  
KW Parkinson's disease; Huntington's disease; immune disorder;  
KW autoimmune disorder; multiple sclerosis; diabetes; allergy.  
XX  
XX  
XX Homo sapiens.  
XX  
XX  
XX WO200218424-A2.  
XX  
XX PD 07-MAR-2002.  
XX  
XX PF 31-AUG-2001; 2001WO-US027093.  
XX  
XX PR 01-SEP-2000; 2000US-00654935.  
XX  
XX (HYSE-) HYSEQ INC.  
XX  
XX Tang YT, Asundi V, Zhou P, Xue AJ, Ren F, Zhang J, Wang J;  
PI Zhao QA, Wang D, Liu C, Drmanac RT, Wehrman T;  
XX  
XX WPI; 2002-583321/62.  
XX  
XX N-PSDB; ABQ93362.  
XX  
XX New polynucleotide and polypeptides, useful for treatment and diagnosis  
PT of Alzheimer's, Parkinson's, Huntington's, amyotrophic lateral  
PT sclerosis, immune deficiencies, cancer, autoimmune disorders, multiple  
PT sclerosis, diabetes and allergies.  
XX  
XX  
XX Claim 20; SEQ ID NO 320; 284pp + Sequence Listing; English.  
XX  
XX The invention relates to an isolated polynucleotide (I) comprising one of  
CC 245 sequences (ABQ93288-ABQ93532). Treating a condition comprising

CC administering to a mammalian subject a composition comprising the protein  
 CC (II) encoded by (I) (ABP62809-ABP63053) or an antibody (III) to (II).  
 CC (I), (II) and (III) are useful for diagnostic evaluation of disorders.  
 CC (I) is useful for gene therapy of diseases and (II) can be used for  
 CC therapeutic treatment. Diseases that may be treated include wound healing  
 CC and tissue repair, burns, central nervous system disorders (e.g.  
 CC Alzheimer's, Parkinson's, Huntington's and amyotrophic lateral  
 CC sclerosis), immune deficiencies, cancer, autoimmune disorders, multiple  
 CC sclerosis, diabetes and allergies. Note: The sequence data for this  
 CC patent did not form part of the printed specification, but was obtained  
 CC in electronic format directly from WIPO at  
 CC ftp.wipo.int/pub/published\_pct\_sequences  
 XX

SO Sequence 1082 AA;

Query Match 58.0%; Score 1364.5; DB 5; Length 1082;

Best Local Similarity 83.2%; Pred. No. 1.1e-137;

Matches 272; Conservative 0; Mismatches 0; Indels 55; Gaps 1;

```

QY 120 KRINQYPSNAEYIASLFPDSLIVKGFNVVSAMALQGPDSRQVYICSNNIQARQQVI 179
DB 1 KRINQYPSNAEYIASLFPDSLIVKGFNVVSAMALQGPDSRQVYICSNNIQARQQVI 60
QY 180 ELARQUNFIPIDGSISSARBIENLPLRLFTLMRGVVAISLATFPFLYSFVDVHPY 239
DB 61 ELARQUNFIPIDGSISSARBIENLPLRLFTLMRGVVAISLATFPFLYSFVDVHPY 120
QY 240 ARNQGSDFYKIPLEIVNKTLPVATTLISLVYIAGLAAAYQLYGTGKRRPPLLETWL 299
DB 121 ARNQGSDFYKIPLEIVNKTLPVATTLISLVYIAGLAAAYQLYGTGKRRPPLLETWL 180
QY 300 QCRKQIGLISFPFAMVHVAYSCLPMRSEERYLFLNMAVQVHANINSEWNEEYVRIEM 359
DB 181 QCRKQIGLISFPFAMVHVAYSCLPMRSEERYLFLNMAVQ----- 220
QY 360 YISFGIMSGLISLAVTSIPSVSNALNMRERFSIQSTLGYALLISTFHVLIYGMKRAF 419
DB 221 -----QSTLGYALLISTFHVLIYGMKRAF 245
QY 420 EEEYRFRYTPPNFVLALVLPISIVIDL 446
DB 246 EEEYRFRYTPPNFVLALVLPISIVIDL 272

```

Search completed: March 1, 2004, 23:33:05  
 Job time : 102 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 1, 2004, 23:31:21 ; Search time 45 Seconds

(without alignments)  
970,466 Million cell updates/sec

Title: US-09-455-486-6

Perfect score: 2351  
Sequence: 1 MESISMGMGSPKSLSETCLPN.....ALVLPSTIVILDLQLCRYPD 454

Scoring table:

BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database :  
1: pir1:\*  
2: pir2:\*  
3: pir3:\*  
4: pir4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	160	6.8	198	2 D95285	conserved hypochet
2	159.5	6.8	211	2 AC2560	hypothetical prote
3	156.5	6.7	239	2 T50571	probable oxidoredu
4	147	6.3	213	2 H69400	conserved hypochet
5	142	6.0	212	2 D69361	conserved hypochet
6	142	6.0	224	2 T10120	F420-dependent NAD
7	133	5.7	223	2 D64487	hypothetical prote
8	121	5.1	191	2 F86826	hypothetical prote
9	120.5	5.1	232	2 A69131	conserved hypochet
10	119.5	5.1	216	2 T00121	hypothetical prote
11	116.5	5.0	242	2 G82642	conserved hypochet
12	116	4.9	198	2 AB3182	conserved hypochet
13	111.5	4.7	695	1 JN0898	foliitropin recept
14	111.5	4.7	1228	1 S59681	hypothetical prote
15	110.5	4.7	222	2 B84410	hypothetical prote
16	110.5	4.7	694	2 J04301	foliitropin recept
17	109	4.6	320	2 T28379	ORF MSV218 hypochet
18	109	4.6	712	2 S50965	probable membrane
19	109	4.6	1184	2 H71436	hypothetical prote
20	109	4.6	1301	2 D85188	disease resistance
21	107.5	4.6	629	2 S60385	foliitropin recept
22	107.5	4.6	695	1 QH0UFT	sodium- and chlori
23	105	4.5	442	2 B64582	hypothetical prote
24	105	4.5	604	2 T31042	chloramphenicol re
25	104	4.4	396	1 C54907	hypothetical prote
26	104	4.4	396	2 G30895	hypothetical prote
27	104	4.4	396	2 H85721	probable resistance
28	103	4.4	465	2 S69915	sodium-phosphate t
29	102.5	4.4	314	2 AB2972	hypothetical prote

30	102.5	4.4	314	2 H98310	probable peptide A
31	102.5	4.4	1242	2 T39453	probable mra stab
32	102	4.3	220	2 AG3547	bicyclicomycin resi
33	102	4.3	574	2 T41068	hypothetical prote
34	101	4.3	320	2 E71139	hypothetical prote
35	101	4.3	348	2 T12284	NADH2 dehydrogenas
36	101	4.3	420	2 F69144	NADH2 dehydrogenas
37	101	4.3	501	2 T02134	O-antigen transpor
38	101	4.3	503	2 C86250	hypothetical prote
39	101	4.3	735	2 A83006	hypothetical prote
40	100.5	4.3	346	2 T11181	NADH2 dehydrogenas
41	100.5	4.3	1780	2 A85045	probable glucan sy
42	100	4.3	395	2 C71219	hypothetical prote
43	100	4.3	476	2 A28439	endonuclease Scat
44	100	4.3	714	2 AF2479	ABC transporter At
45	99.5	4.2	452	2 C71391	NADH2 dehydrogenas

#### ALIGNMENTS

##### RESULT 1

D95285  
conserved hypothetical protein Sma0349 [imported] - Sinorhizobium meliloti (strain 1021)  
C/Species: Sinorhizobium meliloti  
C/Date: 24-Aug-2001 #sequence\_revision 24-Aug-2001 #text\_change 30-Sep-2001  
C/Accession: D95285  
R/Barnett, M.J.; Fisher, R.F.; Jones, T.; Komp, C.; Abola, A.P.; Barloy-Hubler, F.; Bow  
., K.; Kaiman, S.; Keating, D.H.; Palm, C.; Peck, M.C.; Surzycki, R.; Wells, D.H.; Yeh, K.C  
Proc. Natl. Acad. Sci. U.S.A. 98, 9883-9888, 2001  
A/Title: Nucleotide sequence and predicted functions of the entire Sinorhizobium melilo  
A/Reference number: A95262; MUID:21396509; PMID:11481432  
A/Accession: D95285  
A/Status: preliminary  
A/Molecule type: DNA  
A/Residues: 1-198 <KUR>  
A/Cross-references: GB:AE006469; PIDN:AAK64846.1; PID:G14523260; GSPDB:GN00165  
R/Experimental source: strain 1021, megaplasmid pSymA  
R/Galibert, F.; Finan, T.M.; Long, S.R.; Puhler, A.; Abola, P.; Ampe, F.; Barloy-Hubler  
pela, D.; Chain, P.; Cowie, A.; Davis, R.W.; Dreano, S.; Federspiel, N.A.; Fisher, R.F.  
L.; Hyman, R.W.; Jones, T.  
Science 293, 668-672, 2001  
A/Authors: Kahn, D.; Kahn, M.L.; Kaiman, S.; Keating, D.H.; Kiss, E.; Komp, C.; Lelaure,  
heault, P.; Vandenbol, M.; Vorkolter, F.; Weidner, S.; Wells, D.H.; Wong, K.; Yeh, K  
A/Title: The composite genome of the legume symbiont Sinorhizobium meliloti.  
A/Reference number: A96039; MUID:21368234; PMID:11474104  
A/Contents: annotation  
C/Genetics:  
A/Genes: Sma0349  
A/Genome: plasmid

Query Match 6.8%; Score 160; DB 2; Length 198;  
Best local similarity 26.7%; Pred. No. 2.1e-05;  
Matches 55; Conservative 40; Mismatches 83; Indels 28; Gaps 8;

QY	32	TVGVIGSDPAKSLTRILRCGHVIT--GSRNPKAS--EEFPHVVDVTHEDALITKN 87
DB	3	TVAIGAGIGSALERFPAQIPAIANSRGPASISVTDFFGASVAVEIKDAL-QAD 61
QY	88	IIFVAIHRHYTSLMDRLHLV---GKILIDVSNMRLINQV-----ESNAEYASLFP 138
DB	62	VITLAV---PYSDINDIVTVQVSDMGQIVDASNAIDPFAFPRDLGRLSLTEIVSELP 118
QY	139	DSLIVKGFNVVSAMALQGPK--DASROYTCSNNIQARQYIELARQINLPIDLSIS 196
DB	119	GAKVYVAFNTLPAAYLAADPDKGTGSRYLFLSGNSDANRQYAEISSLGFAPVDLGTIA 178
QY	197	SARETENLPRLFTLMRGPVVAISL 222
DB	179	ASGPITQGF-----GRPLVALNL 195

RESULT 2

AC2560  
 hypothetical protein alr8074 [imported] - Nostoc sp. (strain PCC 7120) plasmid pCC7120ga  
 C/Species: Nostoc sp. PCC 7120  
 A/Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120  
 C/Date: 14-Dec-2001 #sequence\_revision 14-Dec-2001 #text\_change 09-Dec-2002  
 C/Accession: AC2560  
 R/Kaneko, T.; Nakamura, Y.; Molk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriuguchi, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, S.  
 DNA Res. 8, 205-213, 2001  
 A/Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Anabaena PCC 7120  
 A/Reference number: AB1807; MUID:21595285; PMID:11759840  
 A/Accession: AC2560  
 A/Status: preliminary  
 A/Molecule type: DNA  
 A/Residues: 1-211 <KOE>  
 A/Cross-references: GB:AP003603; PIDN:BA877404.1; PID:q17134847; GSPDB:GN00182  
 A/Experimental source: strain PCC 7120  
 C/Genetics:  
 A/Gene: alr8074  
 A/Genome: plasmid

Query Match 6.8%; Score 159.5; DB 2; Length 211;  
 Best Local Similarity 24.9%; Pred. No. 2.4e-05;  
 Matches 51; Conservative 43; Mismatches 70; Indels 41; Gaps 8;

QY 33 VGVIGSGDPFKSLTIRLRGCVHVVIGSRNF-----KQSEFFPHVVDVTHEDALTKTN 87  
 DB 3 IAFITIGVGSAALASQSLDITVTIAANSNSDSVKTKALAKYPE-LQVSPQEKIAQAE 61  
 QY 88 IIFVA-----IHREHYTSLMDLRHLVGLKILIDVSNMKNIN-----QYPSNAEYLAFLP 138  
 DB 62 VIFLATPFTANQALAEVGLDLS---GKILVCTNPVGNVLTGLKSGSSSELVQSFVP 117  
 QY 133 DDLIYKGFNVVAMALQLGPKDASROY-----ICSNINQANQVTELRQINF 187  
 DB 118 HAKVVAFT-----YGFENFENHTYPPYGNLKPAMLIAGNDVPAQVVTLCQQLGW 170  
 QY 188 IPIDGLSSASAREIENLPLRLFTLW 212  
 DB 171 EAVDVGNLSMLHLEHMTL---LW 191

# RESULT 3

T50571  
 probable oxidoreductase [imported] - Streptomyces coelicolor  
 C/Species: Streptomyces coelicolor  
 C/Date: 21-Jul-2000 #sequence\_revision 21-Jul-2000 #text\_change 28-Jul-2000  
 C/Accession: T50571  
 R/Redenbach, M.; Kleiser, H.M.; Denapate, D.; Eichner, A.; Cullum, J.; Kinashi, H.; Hopk, M.; Microbiol. 21, 77-96, 1996  
 A/Title: A set of ordered cosmids and a detailed genetic and physical map for the 8 Mb S. coelicolor chromosome  
 A/Reference number: Z20556; MUID:97000351; PMID:8843436  
 A/Accession: T50571  
 A/Status: preliminary; translated from GB/EMBL/DBJ  
 A/Molecule type: DNA  
 A/Residues: 1-239 <RED>  
 A/Cross-references: EMBL:AL133220; PIDN:CAB61708.1  
 A/Experimental source: strain A3(2)  
 C/Genetics:  
 A/Note: SCCT5A.08c  
 C/Superfamily: conserved hypothetical protein MUI501

Query Match 6.7%; Score 156.5; DB 2; Length 239;  
 Best Local Similarity 27.9%; Pred. No. 5e-05;  
 Matches 61; Conservative 41; Mismatches 72; Indels 45; Gaps 10;

QY 25 IDAKRVTVGVV-GSGDFPKSLTIRLRGCVHVVIGSR---NPKASFPFHVDVTHHE 80  
 DB 23 LBDVSGLVGVGCGTGGPGKGLAYRLAKRGKIVIGSRAERAAALAEIGGVGADNA 82  
 QY 81 DALTKNIIFAIHRH-HYTSIMDLRHLVGLKILIDVSNMKNINQ-----YFE-SNAE 131  
 DB 83 ETARRSDVTVAVVPMVDGDKTLESIRAEISGKLVVDVNPVLPFDKKAAYALKEEGSAAE 142

QY 132 YLASLPDSLIVGFFNVVSAMALQ-----LGPKDSROYICSNINQARQCV 178  
 DB 143 QAAALPDSNVAAAFHHLASVILLQDPEIDIDTDVWVLEERRADVAI-----VQA----- 192  
 QY 179 IEARQINFPP-----IDLGSLSAREIENLPLRLFTLWR 213  
 DB 193 --LAGR---LPNGRGVAGLRNAAHQVESLVANLISVNR 226

# RESULT 4

H69400  
 conserved hypothetical protein ARI209 - Archaeoglobus fulgidus  
 C/Species: Archaeoglobus fulgidus  
 C/Date: 05-Dec-1997 #sequence\_revision 05-Dec-1997 #text\_change 29-Sep-1999  
 C/Accession: H69400  
 R/Klenk, H.P.; Clayton, R.A.; Tomb, J.F.; White, O.; Nelson, K.E.; Ketchum, K.A.; Dodson, J.; Fleischmann, R.D.; Quackenbush, J.; Lee, N.H.; Sutton, G.G.; Gill, S.; Kirkness, E.F.; Glodek, A.; Zhou, L.; Overbeek, R.; Gocayne, J.D.; Weidman, J.F.; McDonald, L.  
 Nature 390, 364-370, 1997  
 A/Authors: Overbeek, T.; Cotton, M.D.; Spriggs, T.; Artiach, P.; Kaine, B.P.; Sykes, S.; Smith, H.O.; Woese, C.R.; Venter, J.C.  
 A/Title: The complete genome sequence of the hyperthermophilic, sulfate-reducing archaeon Archaeoglobus fulgidus  
 A/Reference number: A69250; MUID:98049343; PMID:9389475  
 A/Accession: H69400  
 A/Status: preliminary; nucleic acid sequence not shown; translation not shown  
 A/Molecule type: DNA  
 A/Residues: 1-213 <KOE>  
 A/Cross-references: GB:AE001021; GB:AE000782; NID:G2689344; PIDN:AAB90038.1; PID:G26493  
 C/Superfamily: conserved hypothetical protein MUI501

Query Match 6.3%; Score 147; DB 2; Length 213;  
 Best Local Similarity 27.5%; Pred. No. 0.00024;  
 Matches 56; Conservative 44; Mismatches 68; Indels 36; Gaps 11;

QY 37 GSGDFPKSLTIRLRGCVHVVIGSRNF-----KQSEFFPHVVDVTHEDALTKTN 87  
 DB 8 GTGNLGGALRKGKGLYEIVGSRKLEKAEKIASDYLKVGDS-IGKRNEDAAETCD- 66  
 QY 89 IFAIHRHYTSLMD-----LRHLVGLKILID-----VSNMKNINQPE-SNAE 131  
 DB 67 --VAV---FTLPWFAPFTAEMLKQAGKAVVISPVMKVGDNF-YVAREEGSAAE 119  
 QY 132 YLASLPDSLIVGFFNVVSAMALQLGPKDASROYICSNINQARQCVIEARQINF-PI 190  
 DB 120 KLASVLESVVAAHYSIPARRPANGGEFFEMDVPIGDS-GAKEVVVDLTKISGLRAL 178  
 QY 191 DLGSLSAREIENL-PLRLFTLWR 213  
 DB 179 DAGGLSNAHLVBSLTPLLIINVKR 202

# RESULT 5

D69361  
 conserved hypothetical protein AF0892 - Archaeoglobus fulgidus  
 C/Species: Archaeoglobus fulgidus  
 C/Date: 05-Dec-1997 #sequence\_revision 05-Dec-1997 #text\_change 29-Sep-1999  
 C/Accession: D69361  
 R/Klenk, H.P.; Clayton, R.A.; Tomb, J.F.; White, O.; Nelson, K.E.; Ketchum, K.A.; Dodson, J.; Fleischmann, R.D.; Quackenbush, J.; Lee, N.H.; Sutton, G.G.; Gill, S.; Kirkness, E.F.; Glodek, A.; Zhou, L.; Overbeek, R.; Gocayne, J.D.; Weidman, J.F.; McDonald, L.  
 Nature 390, 364-370, 1997  
 A/Authors: Overbeek, T.; Cotton, M.D.; Spriggs, T.; Artiach, P.; Kaine, B.P.; Sykes, S.; Smith, H.O.; Woese, C.R.; Venter, J.C.  
 A/Title: The complete genome sequence of the hyperthermophilic, sulfate-reducing archaeon Archaeoglobus fulgidus  
 A/Reference number: A69250; MUID:98049343; PMID:9389475  
 A/Accession: D69361  
 A/Status: preliminary; nucleic acid sequence not shown; translation not shown  
 A/Molecule type: DNA  
 A/Residues: 1-212 <KOE>  
 A/Cross-references: GB:AE001042; GB:AE000782; NID:G2689365; PIDN:AAB900348.1; PID:G26493  
 C/Superfamily: conserved hypothetical protein MUI501

Query Match 6.0%; Score 142; DB 2; Length 212;  
 Best Local Similarity 25.4%; Pred. No. 0.0006;  
 Matches 49; Conservative 44; Mismatches 64; Indels 16; Gaps 6;

QY 37 GSGDPAKSLITRLIRCGYHVIIGSSNPK---FASEFPFHVDVT---HHEDALTQNI 88  
 DB 8 GTGNLKGKGLALRLATLGEIIVGSSRREKAEKAEYRIADASITGKVEDAAEACDI 67  
 QY 89 IFVAIHREH-YTSLMDRLRLVGLKILIDV-----LIDVSNMRLNQY--PESNAEYIASLFPDSL 141  
 DB 68 AVLTTPWEHAIDTARDLKNILREKIVSPLFVSGANGFTYSSERSAAETVAEVLSEK 127  
 QY 142 IYKGFNVASAMALQGPCKDASROYVICSNNIQARQVIELARQNLFI-PIDIGLSISARE 200  
 DB 128 VYSALATTIPAAEFANLDEKFDVPCGDDDSKVKVMSLISEIDGLPLDAGPLSNRL 187  
 QY 201 IENLPLRLFTLMR 213  
 DB 188 VESLPLRLINMR 200

RESULT 6  
 T10120  
 F420-dependent NADP reductase (EC 1.6.8.-) - Methanobacterium thermoautotrophicum (strain  
 C/Specties: Methanobacterium thermoautotrophicum  
 C/Date: 16-Jul-1999 #sequence\_revision 16-Jul-1999 #text\_change 20-Jun-2000  
 C/Accession: T10120  
 R/Berk, H.; Thauer, R.K.  
 FEBS Lett. 438, 124-126, 1998  
 A/Title: F420H2:NADP oxidoreductase from Methanobacterium thermoautotrophicum: identifi  
 A/Reference number: Z16959; M01D:19037734; PMID:9821972  
 A/Accession: T10120  
 A/Status: preliminary; translated from GB/EMBL/DBJ  
 A/Molecule type: DNA  
 A/Residues: 1-224 <BER>  
 A/Cross-references: EMBL.Y17210  
 A/Experimental source: strain Marburg  
 A/Suprafamily: conserved hypothetical protein M01501  
 C/Keywords: oxidoreductase

Query Match 6.0%; Score 142; DB 2; Length 224;  
 Best Local Similarity 29.2%; Pred. No. 0.00064;  
 Matches 59; Conservative 29; Mismatches 84; Indels 30; Gaps 8;

QY 37 GSGDPAKSLITRLIRCGYHVIIGSSNPKFASEFPFHVDVTHHEDALTQNI-----TN----- 87  
 DB 8 GTGDOGGLGIALRLALAGEEVIGSRDAKAVSAQVLEIERDDLKYKATNAEABEA 67  
 QY 88 ---IIFVAIHREHYTSLMDRLRLVGLKILIDV-----SNNMRINQYPESSAEYLA 134  
 DB 68 EVALLTVPLOQOMAT-LGSVKEATKXKVLDTVPIDSCLOGSAVRYTLDWGSAAERPA 126  
 QY 135 SLFPD--SLIVKGFNVASAMALQ--LGPCKDASROYVICSNNIQARQVIELARQNLFI-P 189  
 DB 127 RLIEDQGRVAAAFNNISASALDLITGPDC--DCLINSDHRDLADLASAEKIDGVA 184  
 QY 190 IDLGSLSAREIENLPLRLFTL 211  
 DB 185 IDCGGLENARVIEKITPLINL 206

RESULT 7  
 D64487  
 Hypothetical protein M01501 - Methanococcus jannaschii  
 C/Specties: Methanococcus jannaschii  
 C/Date: 13-Sep-1996 #sequence\_revision 13-Sep-1996 #text\_change 21-Jul-2000  
 C/Accession: D64487  
 R/Bult, C.J.; White, O.; Olsen, G.T.; Zhou, L.; Fleischmann, R.D.; Sutton, G.G.; Blake,  
 Reich, C.T.; Overbeek, R.; Kirkness, E.F.; Weissbrock, K.G.; Merrick, J.M.; Glodek, A.;  
 Iron, J.D.; Sadow, P.W.; Hanna, M.C.; Cotton, M.D.; Roberts, K.M.; Hurst, M.A.  
 Science 273, 1058-1073, 1996  
 A/Authors: Kaine, B.P.; Borodovsky, M.; Klenk, H.P.; Fraser, C.M.; Smith, H.O.; Woese, C  
 A/Title: Complete genome sequence of the methanogenic archaeon, Methanococcus jannaschii

A/Reference number: A64300; M01D:96337999; PMID:8688087  
 A/Accession: D64487  
 A/Status: preliminary; nucleic acid sequence not shown; translation not shown  
 A/Molecule type: DNA  
 A/Residues: 1-223 <BUL>  
 A/Cross-references: GB:U67591; GB:L77117; NID:G286422; PIDN:AA89514.1; PID:G1500389;  
 C/Genetics:  
 A/Map position: REV1473617-1472946  
 C/Suprafamily: conserved hypothetical protein M01501

Query Match 5.7%; Score 133; DB 2; Length 223;  
 Best Local Similarity 21.9%; Pred. No. 0.0033;  
 Matches 46; Conservative 52; Mismatches 78; Indels 34; Gaps 7;

QY 37 GSGDPAKSLITRLIRCGYHVIIGSSNPKFASEFPFHVDV-----THHEDALT 84  
 DB 8 GTGDOGGLGIALRLAK-NKTIISRKKEKAEKAEKARETLKORGIEADITGLENDAK 66  
 QY 85 KTNIIFVAIHREH-YTSLMDRLRLVGLKILIDV-----SNNMRINQYPE--SNAEYL 133  
 DB 67 EGDVILSLPEYETLSTINGQLKEELKGIYVSGVPLATAGDKPRLPLPPDGSAEMV 126  
 QY 134 ASLFPDSLIVKGFNVASAMALQGPCKDASROYVICSNNIQARQVIELARQNLFI-PIDL 192  
 DB 127 QNVLKESKVVSAFQNVCHAVLEDLNVPDCDILVCGNDEAKKVVIDLANOIGVRAIDC 186  
 QY 193 GSLSAREIENLPLRLFTLMRGPVVAISL 222  
 DB 187 GNLKSKRIITAIT-----DLIGLNI 207

RESULT 8  
 F68826  
 Hypothetical protein yf6e [imported] - Lactococcus lactis subsp. lactis (strain IL1403)  
 C/Specties: Lactococcus lactis subsp. lactis  
 C/Date: 23-Mar-2001 #sequence\_revision 23-Mar-2001 #text\_change 03-Aug-2001  
 C/Accession: F68826  
 R/Bolotin, A.; Manger, P.; Manger, S.; Jallion, O.; Malarme, K.; Weisenbach, J.; Ehrh  
 Genome Res. 11, 731-753, 2001  
 A/Title: The complete genome sequence of the lactic acid bacterium Lactococcus lactis s  
 A/Reference number: A86625; M01D:21235186; PMID:11337471  
 A/Accession: F68826  
 A/Status: preliminary  
 A/Molecule type: DNA  
 A/Residues: 1-191 <STO>  
 A/Cross-references: GB:AE005176; PID:G12724621; PIDN:AAK05712.1; GSPDB:GND0146  
 A/Experimental source: strain IL1403  
 C/Genetics:  
 A/Gene: yf6e

Query Match 5.1%; Score 121; DB 2; Length 191;  
 Best Local Similarity 22.2%; Pred. No. 0.024;  
 Matches 47; Conservative 33; Mismatches 74; Indels 56; Gaps 7;

QY 32 TVGVISGDPKSLITRLIRCGYHVIIGSSNPKFASEFPFHVDVTHHEDALTQNI----- 87  
 DB 3 TISIFGKXGKXKI-----GNFSSVYKVVYILSNSKTELGET 42  
 QY 88 ---IIFVAIHREHYTSLMDRLRLVGLKILIDVSNMRINQYPE-----SNAEYL 133  
 DB 43 VVLAAPVYVLAGIIGYSTD-----LQKIIIDITNVDFTTSPSLVPSYSAALI 95  
 QY 134 ASLFPDSLIVKGFNVASAMALQGPCKDASRQ---YVCSNNIQARQVIELARQNLFI-P 189  
 DB 96 AKQLPNSMIVKAN--TTFSDTLATKRVANHEQTIVTLASDQAEKRTIIKALENSGLSL 153  
 QY 190 IDLGSLSAREIENLPLRLFTL-----WRG 214  
 DB 154 LDAGSLKRAELAEIQLITTLAASKISWDG 185

RESULT 9  
 A69131

conserved hypothetical protein MTH248 - Methanobacterium thermoautotrophicum (strain Del  
C/Species: Methanobacterium thermoautotrophicum  
C/Date: 05-Dec-1997 #sequence\_revision 05-Dec-1997 #text\_change 29-Sep-1999  
C/Accession: A69131  
R/Smith, D.R.; Doucette-Stamm, L.A.; DeJonghery, C.; Lee, H.; Dubois, J.; Aldredge, T.;  
Qiu, D.; Spadefora, R.; Vitacite, R.; Wang, Y.; Wierzbowski, J.; Gibson, R.; Jivanti, N.  
K.; Church, G.M.; Daniels, C.J.; Mao, J.; Rice, P.; Noelling, J.; Reeve, J.N.  
J. Bacteriol. 179, 7135-7155, 1997  
A/Title: Complete genome sequence of Methanobacterium thermoautotrophicum Delta H: funct  
A/Reference number: A69000; MUID:98037514; PMID:9371463  
A/Accession: A69131  
A/Status: preliminary; nucleic acid sequence not shown; translation not shown  
A/Molecule type: DNA  
A/Residues: 1-232 <MTH-  
A/Cross-references: GB:AE009811; GB:AE006666; NID:g2621287; PIDN:AA84754.1; PID:g262128  
A/Experimental source: strain Delta H  
C/Genetics:  
A/Gene: MTH248  
A/Start codon: TTG  
C/Superfamily: conserved hypothetical protein MJI501

Query Match 5.1%; Score 120.5; DB 2; Length 232;  
Best Local Similarity 23.8%; Pred. No. 0.034; Indels 25; Gaps 6;  
Matches 49; Conservative 42; Mismatches 90; Indels 25; Gaps 6;

QY 30 KTVGVIGSGDFPAKSLTIRLCGYHVIGSRNPKFASEFPFVVDVTHHE----- 80  
DB 8 KIAV-IGGTGGDLALRFNVAAGEVITIGSRDSEKAKSKYLEINGRDISEGATN 66  
QY 81 -DALTKNIIFAVH-RHYTSLMDRLVGLKIIIVS-----NNKINQYPSN 129  
DB 67 PDAAASAVVVLTVLPQOMVTLASIRQVDRKVIDATVDSICIGSAARYIDLEGS 126  
QY 130 AEYLASF---PDSLVIGFNVVSAMALQGPDKASRQVYCSNNIQARQVIELARQIN 186  
DB 127 AAFRAARFLREGRTVAAAFNNISALLEVSEPDCCCLVASHRDLVFAAELEKID 186  
QY 187 FI-PIDGLSLGSAEIEMLPRLFTL 211  
DB 187 GVRAIECGLEVARIEKITEPLINL 212

RESULT 10  
T00121  
hypothetical protein 8 - Leptospira interrogans  
C/Species: Leptospira interrogans  
C/Date: 22-Jan-1999 #sequence\_revision 22-Jan-1999 #text\_change 21-Jul-2000  
C/Accession: T00121  
R/Takahashi, Y.; Akase, K.; Hirano, H.; Fukunaga, M.  
Gene 215, 37-45, 1998  
A/Title: Physical and genetic maps of the Leptospira interrogans serovar icterohaemorrh  
A/Reference number: Z14115; MUID:98332717; PMID:9666070  
A/Accession: T00121  
A/Status: preliminary; translated from GB/EMBL/DBJ  
A/Molecule type: DNA  
A/Residues: 1-216 <TAK-  
A/Cross-references: EMBL:AB010203; NID:g2780763; PIDN:BA24368.1; PID:g2780765  
A/Experimental source: strain ictero No.1; substrain icterohaemorrhagiae

Query Match 5.1%; Score 119.5; DB 2; Length 216;  
Best Local Similarity 23.7%; Pred. No. 0.037; Indels 23; Gaps 6;  
Matches 47; Conservative 37; Mismatches 91; Indels 23; Gaps 6;

QY 33 VAVTSSGDFPAKSLTIRLCGYHVIGSRNPKFASEFPFV---VDVTHEDALTKNII 89  
DB 6 IGLTSGVIGVQTLANGFAGVETIGTRFGLKDWLAKAGAGAGSFGSEANFGEIT 65  
QY 90 FVAIHRHYTSLMDRLH-LVQKILIDVSNMR-----INQYPSNAEYASL 136  
DB 66 VLCSKGSVASVLTLSGIDSLNGKTIIDTNPDISIPQNGVNLFPISYNSLMEKQK 125  
QY 137 PDSLVIGFNVVSAMALQGP---KDSARQYVIGSNNIQARQVIELARQINFIPI 194

DB 126 AFAANVVKCRSSVGS-GLVNPQQLKGRKPSMFIQGNDSKKQIKIILDTFGWIDEDMK 184  
QY 195 LSSAREIENLPRLFTLW 212  
DB 185 VEARAIEPLCI---LW 198

RESULT 11  
G82642  
conserved hypothetical protein XF1737 [imported] - Xylella fastidiosa (strain 9a5c)  
C/Species: Xylella fastidiosa  
C/Date: 18-Aug-2000 #sequence\_revision 20-Aug-2000 #text\_change 20-Aug-2000  
C/Accession: G82642  
R/Anonymous, The Xylella fastidiosa Consortium of the Organization for Nucleotide Sequ  
Nature 406, 151-157, 2000  
A/Title: The genome sequence of the plant pathogen Xylella fastidiosa.  
A/Reference number: A82515; MUID:20365717; PMID:10910347  
A/Note: for a complete list of authors see reference number A59328 below  
A/Accession: G82642  
A/Status: preliminary  
A/Molecule type: DNA  
A/Residues: 1-242 <SIM-  
A/Cross-references: GB:AE003997; GB:AE003849; NID:g9106805; PIDN:AAF4546.1; GSPDB:GN00  
R/Simpson, A.J.G.; Reinach, F.C.; Arruda, P.; Abreu, F.A.; Acencio, M.; Alvarenga, R.;  
Briones, M.R.S.; Bueno, M.R.P.; Camargo, A.A.; Camargo, L.E.A.; Carreiro, D.M.; Carrer,  
as-Neto, E.; Docena, C.; El-Dorri, H.; Facincani, A.B.; Ferreira, A.J.S.  
submitted to Genbank, June 2000  
A/Authors: Ferreira, V.C.A.; Ferro, J.A.; Fraga, J.S.; Franca, S.C.; Franco, M.C.; Froh  
J.D.; Junqueira, M.L.; Kemper, E.L.; Kitzajima, J.P.; Krieger, J.E.; Kuramae, E.B.; Laig  
chado, M.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; Marino, C.L.; Marques, M.V.; Martins,  
A. Authors: Martins, E.M.F.; Matsukuma, A.Y.; Menck, C.F.M.; Miracca, E.C.; Miyaki, C.Y.  
F.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Palmieri, D.  
Rodrigues, V.; Rosa, A.U. de M.; de Rosa, V.E.; de Sa, R.G.; Sanello, R.V.; Sawa  
A. Authors: da Silva, A.C.R.; da Silva, F.R.; da Silva, A.M.; Silva Jr., M.A.; da Silve  
M.; Tshako, M.H.; Vallada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L.;  
A/Reference number: A59328  
A/Contents: annotation  
C/Genetics:  
A/Gene: XF1737

Query Match 5.0%; Score 116.5; DB 2; Length 242;  
Best Local Similarity 24.0%; Pred. No. 0.075; Indels 21; Gaps 7;  
Matches 46; Conservative 40; Mismatches 85; Indels 21; Gaps 7;

QY 28 ARKVTYVIGSGDFPAKSLTIRLCGYHVIGSRNPKFASEFPFVVDVTHHEDEL 83  
DB 30 AAFMRIGVIGAGSLGTGRLVWKGHEVMPFSRNPDKLEAARLEPR-ASVGGPLAT 86  
QY 84 TKTNIIFVAIHRHYTSLM-DIRHLVGLKILIDVSNMRINO---YPSN---AEYLAS 135  
DB 89 EFGTVLLAVPPEALPQVDRDLSAYRGKIVDSITPFGASADYVREARLEGVQTVK 148  
QY 136 LFPDSLVIGFNVVSAMALQGPDKASRQ-----YICSNNIQARQVIELARQINFIPI 190  
DB 149 YMGARLVAFSAVDATVH---TSASRGRGRIQNPILASDDBAKVAAGLVYRDAGCPV 205  
QY 191 DUGLSSSAREIE 202  
DB 206 IVGNLAAASFPQ 217

RESULT 12  
AB3182  
conserved hypothetical protein Abus183 [imported] - Agrobacterium tumefaciens (strain C  
C/Species: Agrobacterium tumefaciens  
C/Date: 11-Jan-2002 #sequence\_revision 11-Jan-2002 #text\_change 18-Nov-2002  
C/Accession: AB3182  
R/Wood, D.W.; Sebval, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Moo  
erge, G.; Giller, W.; Grant, C.; Guenther, D.; Kutyavain, T.; Levy, R.; Li, M.; McClell  
i Karp, F.; Romero, P.; Zhang, S.  
Science 294, 2317-2323, 2001



A:Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm, J.E.W.  
 A:Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.  
 A:Reference number: AB2577; MUID:21608550; PMID:11743193  
 A:Accession: AB3182  
 A:Status: Preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-198 <KDR>  
 A:Cross-references: GS:AE006687; PIDN:AA45872.1; PID:G17743615; GSPDB:GN00188  
 A:Experimental source: strain C58 (Dupon)  
 C:Genetics:  
 A:Gene: Atub183  
 A:Genome: plasmid

Query Match 4.9%; Score 116; DB 2; Length 198;  
 Best Local Similarity 22.7%; Pred. No. 0.063;  
 Matches 50; Conservative 41; Mismatches 81; Indels 48; Gaps 10;

QY 31 VTGVIGSGDPAKSLTIRLRIGYHVI-GSRNPKFASFPFHV---VDVTHEDALT 86  
 DB 1 MTVGIGAGNIGAFATLKGAGIEAVIANGSGESLIALVSKLSTIRAGSVPEAAQA 60  
 QY 87 NIIVAIHREHYTSLMD-----LRHLVG-KILIDVSNKRIINQY-----PESNAEYLA 134  
 DB 61 -IVVAIV-----WSKIPGALAGLNFGRIVDANNISIEAPLYRPADLGRSTDI 112  
 QY 135 SLFPELSLVKGFNVSAWALQGPDKAS-----ROYICGNINQARQVETELAROL 185  
 DB 113 ALVGAARVYKAFN-----HLTRKQSGDPHSGGRVLYFSQSDMKALAEVGAIDRI 165  
 QY 186 NFIPIDGSLSSAREIENLPRLFTLMRGPVVAISLATF 225  
 DB 166 GFPGIDIGGLPVGSQLQFP-----GGPLPALNLVKE 197

## RESULT 13

UN0898  
 Follitropin receptor precursor - crab-eating macaque  
 N:Alternate names: follicle-stimulating hormone receptor (FSHR)  
 C/Species: Macaca fascicularis (crab-eating macaque)  
 C/Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 10-Sep-1999  
 C/Accession: UN0898; S36452  
 R:Gromoll, J.; Dankbar, B.; Sharma, R.S.; Nieschlag, E.  
 Biotech. Biophys. Res. Commun. 196, 1066-1072, 1993  
 A:Title: Molecular cloning of the testicular follicle stimulating hormone receptor of the  
 A:Reference number: UN0898; MUID:94071854; PMID:7504463  
 A:Accession: UN0898  
 A:Molecule type: mRNA  
 A:Residues: 1-695 <GRO>  
 A:Cross-references: EMBL:X74454; NID:9396801; PIDN:CA452463.1; PID:9396802  
 A>Note: The authors translated the codon AGT for residue 488 as Arg  
 C/Function:  
 A:Description: receptor that mediates the biochemical effects of follitropin  
 C:Superfamily: glycoprotein hormone receptor; leucine-rich alpha-2-glycoprotein repeat  
 C/Keywords: G protein-coupled receptor; glycoprotein; hormone receptor; phosphoprotein;  
 F.1-17/Domain: signal sequence #status predicted <SIG>  
 F.18-695/Product: follitropin receptor #status predicted <PRH>  
 F.71-95/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR2>  
 F.96-120/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR3>  
 F.121-145/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR4>  
 F.146-169/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR5>  
 F.172-193/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR6>  
 F.194-218/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR7>  
 F.367-387/Domain: transmembrane #status predicted <TM1>  
 F.399-421/Domain: transmembrane #status predicted <TM2>  
 F.444-465/Domain: transmembrane #status predicted <TM3>  
 F.486-508/Domain: transmembrane #status predicted <TM4>  
 F.529-550/Domain: transmembrane #status predicted <TM5>  
 F.574-597/Domain: transmembrane #status predicted <TM6>  
 F.609-630/Domain: transmembrane #status predicted <TM7>  
 F.719-793,293,318/Binding site: carbohydrate (Asn) (covalent) #status predicted  
 F.755/Binding site: phosphate (Thr) (covalent) (by protein kinase C) #status predicted  
 F.596/Binding site: phosphate (Ser) (covalent) (by protein kinase C) #status predicted

Query Match 4.7%; Score 111.5; DB 1; Length 695;  
 Best Local Similarity 17.0%; Pred. No. 0.71;  
 Matches 84; Conservative 75; Mismatches 151; Indels 183; Gaps 16;

QY 79 HEDALKTNIIFVAIHRHYTSLMDLRLHVG-----KILIDVSNMNR 121  
 DB 98 HETRIKANKNL-LYNPEAFQVLPRLRIILLINIGIKHLPVHKHSRQVLLDIQDNIN 156  
 QY 122 INQYPSNAEYLAFLPDSLVKGFNVSAWALQGPDKASROYICGNINQARQVETEL 181  
 DB 157 IH-----TIERNFVGLSFESEVILW-----NKGQIEINCA----- 189  
 QY 182 AQLNPIPIDGSLSSAREIENLPRLFTLMRGPVVAISLATFFLVSPFVDVHPYAR 241  
 DB 190 ---FNGTQDELINLNNLLEELPNDVFAGSGPVILLISRTIRHSLPSYLENLKURA 246  
 QY 242 NOQSPFYKPIRIEIVKTYPIVAITLSTLV----- 271  
 DB 247 RSTYMLKTLF-----SLKLVALMEASLTYPEHCAPANWRQISELHPICNKSILRQEV 301  
 QY 272 -----LAGLAAAYQLYGTGKRRF-----PMLFTWLQCRKQIG 306  
 DB 302 DYMOTQGRGSSLAEDNESYSRGPDMTYAEFDYDLCEVVDVTSPPKDAFNPCEIDIG 361  
 QY 307 -----LSFPFAMVVA-----YSLCLPMRSEERYLFLNNAVQOVHANIENS 348  
 DB 362 VNIKRLVIFISILAITNGIIVLVLTISQYLTVP-----RLNCLNLAFAF----- 408  
 QY 349 KNEEVEVRIEYISFGINSGLSLAIVTSIPSVAALNMR-----PSFIQSTIG 399  
 DB 409 -----LCIGIYLLIASVDIHTKSGYHVAIDMOTGACDGAAGFTVASELS 456  
 QY 400 ---VYALLISTFH-----VLGYKRAPEEZYRYRFPNPNV-LAL 436  
 DB 457 VTTLTATLTERHATTTTHAMQDCKVYHRAASVWGMIFAPAALLFPIFGISYMKVSI 516  
 QY 437 VLPSTVIIDLLQL 449  
 DB 517 CLPMDIDSPLSQL 529

## RESULT 14

S59681  
 Probable membrane protein YP012W - yeast (Saccharomyces cerevisiae)  
 N:Alternate names: hypothetical protein YPA5W; hypothetical protein YP8132.01  
 C/Species: Saccharomyces cerevisiae  
 C/Date: 13-Jan-1996 #sequence\_revision 01-Mar-1996 #text\_change 19-Apr-2002  
 C/Accession: S59681; S52519  
 R/Hall, U.; Ahmed, A.; Bussey, H.; Fortin, N.; Friesen, J.D.; Storms, R.K.; Vo, D.H.; W.  
 submitted to the EMBL Data Library, August 1995  
 A:Description: The sequence of Saccharomyces cerevisiae chromosome XVI left arm.  
 A:Reference number: S59677  
 A:Accession: S59681  
 A:Molecule type: DNA  
 A:Residues: 1-1228 <HAL>  
 A:Cross-references: EMBL:U33335; NID:9365076; PID:9365081; MIPS:YP012W  
 R/Baddock, K.; Churcher, C.  
 submitted to the EMBL Data Library, February 1995  
 A:Reference number: S52519  
 A:Accession: S52519  
 A:Molecule type: DNA  
 A:Residues: 220-1228 <BAD>  
 A:Cross-references: EMBL:248483; NID:9683777; PID:9683778; MIPS:YP012W  
 A:Experimental source: strain AB972  
 C/Genetics:  
 A:Gene: SGD:RP12  
 A:Cross-references: SGD:S0005933  
 A:Map position: 16L  
 C/Keywords: transmembrane protein  
 F.125-141/Domain: transmembrane #status predicted <TM1>  
 F.478-494/Domain: transmembrane #status predicted <TM2>  
 F.720-736/Domain: transmembrane #status predicted <TM3>

Query Match 4.7%; Score 111.5; DB 2; Length 1228;  
 Best Local Similarity 19.9%; Pred. No. 1.5;  
 Matches 111; Conservative 71; Mismatches 152; Indels 225; Gaps 28;

```

QY 11 KSLSECLPNGINGIKDKARK-----VTVGIGSGDFAKSLTITLI 50
DB 292 KMAETITSSGLAENKYLAVLDTIFALRKSNDITLLTSMINAVIKGSTAHTHQPLRL 351
QY 51 R-----CGYHVIGSRNPKF--ASEFFPHVDVTHEDAL----- 83
DB 352 RKIPGVFHMCTY--LASETEPVQAPASQCLISLISESVKDDLTYPSVDEKFKAVD 408
QY 84 -----TKTNIFVAIRREHYTSLMDLRLVGLKILLIDVSNMRLNQVPSNAEYLASTLF 137
DB 409 EITISQIAKTFIDPLSRISHCS-----REIL--KILVAAPKFRY--RSNPFRLKSL 456
QY 138 PSLIYKGVNVSAMALQGPXDASRQVYICSNNTQARQOYIELARQINF1----- 188
DB 457 -----KIVDTWEVN-----BEQFMDLNEIEILVIGASISAMG 488
QY 189 -----PIDGSSAREIENLPLRLFTLRGPV-----VAISLAPFLYSFVRDYI 236
DB 489 PEMILAEADPLNDPESER-----PGR--AWLLPLINDYTKANLATP-----QNEL 533
QY 237 HPYARNQSGDFYKIPLEIVN-----KTLF-----IVAITLTL 269
DB 534 APYIKSFQSKFKVPESIQLRVFCQIVDQIWSTLPRCELPMDLRBSPTDEFASELSL 593
QY 270 VTL-----AGLLAAAYQLYGTG-----YRRFP-----PWLETWLQCRK 303
DB 594 LYSEVELRTTICHLKVALESNVSVAEBSSSHVLLQRFPISEAQKNIETST-----K 648
QY 304 QGLSLFFPAMVAVASLCLPWRSS-----ERYL-----FLNAYQOVHANIENSWNE 351
DB 649 STNLAVLEFN-----YTQTPNARSYLETIDQYLRKSKEDLEKTPNVCGLKSNME 704
QY 352 EEVWRLEMYISFGINSGLLSLL--AVTISPSVSNALNMRFEFTQSTLGYVALLISTFH 409
DB 705 EESGNNKKEKKPOLATLTLIDLICMITYLPIVSSYSALEFMSFS-----LTVNSAD 754
QY 410 VLIYGWRAFEERYEYFYT 428
DB 755 ALIQ--KRA-----YRIIT 766

```

## RESULT 15

B84410  
 hypothetical protein Vng2607c [imported] - Halobacterium sp. NRC-1  
 C/Species: Halobacterium sp. NRC-1  
 C/Date: 02-Feb-2001 #sequence\_revision 02-Feb-2001 #text\_change 16-Feb-2001  
 C/Accession: B84410  
 R/By: W.V.; Kennedy, S.P.; Mahatas, G.G.; Bergquist, B.; Pan, M.; Shukla, H.D.; Lasky, S.; Leithauer, B.; Keller, K.; Cruz, R.; Danson, M.J.; Hough, D.W.; Maddocks, D.G.; Jabil Jung, K.H.; Alam, M.; Freitas, T.  
 Proc. Natl. Acad. Sci. U.S.A. 97, 12176-12181, 2000  
 A/Authors: Hou, S.; Daniels, C.J.; Dennis, P.P.; Omer, A.D.; Ehardt, H.; Lowe, T.M.; Li  
 A/Title: Genome sequence of Halobacterium species NRC-1.  
 A/Reference number: A84160; M01D:20504483; PMID:11016950  
 A/Accession: B84410  
 A/Status: preliminary  
 A/Molecule type: DNA  
 A/Residues: 1-222 <8TO>  
 A/Cross-references: GB:AE004437; NID:gi10561987; PIDN:AA020646.1; GSPDB:GN00138  
 C/Genetics:  
 C/gene: VNG2607C  
 C/superfamily: conserved hypothetical protein MJ1501

Query Match 4.7%; Score 110.5; DB 2; Length 222;  
 Best Local Similarity 24.3%; Pred. No. 0.2;  
 Matches 49; Conservative 34; Mismatches 94; Indels 25; Gaps 6;

```

DB 8 GTDIDAGLRLRWATDSDDHDIVIGSRDPEKARETTAAAYEDTLADQGVDRKLTGFANEMAA 67
QY 84 TKTNIFVAIRREHYTSLW-----DLRLVGLKILLIDVSNMRLNQVPSNAE--YL 133
DB 68 DRADVIVVAAPAYHVTIDVMAQVADRDLADPTLVISPAVGIAAGBGLHYNPPSAGSVTALV 127
QY 134 ASLEFPDSL-IVKFPNVVSAMALQGPXDASRQVYICSNNTQARQOYIELARQINF1-PID 191
DB 128 ADAAPDGVAVGAFHNLADRLADLTDELADPTLVVGNBGAETRVAEILDITGLRALD 187
QY 192 LGSLSAREIENLPLRLFTLMR 213
DB 188 AGPVENAAEYESLTPILINLAR 209

```

Search completed: March 2, 2004, 06:12:46  
 Job time : 51 secs

QY 37 GSGDPAKSLTIRLCGHH-VVIGSRNPKFASEFFPHVVDVT-----HEDAL 83



```

; APPLICANT: Lal, Preeti
; APPLICANT: Guegler, Karl J.
; APPLICANT: Corley, Neil C.
; TITLE OF INVENTION: PROSTATE GROWTH-ASSOCIATED MEMBRANE PROTEINS
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
; STREET: 3174 PORTER DRIVE
; CITY: PALO ALTO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/083,521
; FILING DATE: Herewith
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: CERONE, MICHAEL C.
; REGISTRATION NUMBER: 39,132
; REFERENCE/DOCKET NUMBER: PF-0527 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650) 855-0555
; TELEFAX: (650) 845-4166
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 141 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: PROSIT0110
; CLONE: 1691243
;
US-09-083-521-1
Query Match 31.3%; Score 736; DB 3; Length 141;
Best Local Similarity 100.0%; Pred. No. 2,2e-69;
Matches 141; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 314 MHVAVSLCLPMRBSRYFLINMAVOOVANINSENNBEEVRIEMTISFGIMSLGLLSTL 373
DB 1 MHVAVSLCLPMRBSRYFLINMAVOOVANINSENNBEEVRIEMTISFGIMSLGLLSTL 60
QY 374 LWTSTPSVSNLNMWEEPSFIQSTGYVALLISTFHVLLYGMKRAPEEERYAFYTPNPFV 433
DB 61 LWTSTPSVSNLNMWEEPSFIQSTGYVALLISTFHVLLYGMKRAPEEERYAFYTPNPFV 120
QY 434 LALVLPISIVLIDLQLCRYPD 454
DB 121 LALVLPISIVLIDLQLCRYPD 141

RESULT 3
US-09-323-873A-2
; Sequence 2, Application US/09323873A
; Patent No. 6329503
; GENERAL INFORMATION:
; APPLICANT: Daniel E. Afar
; APPLICANT: Rene S. Hubert
; APPLICANT: Kahan Leong
; APPLICANT: Arthur B. Raitano
; APPLICANT: Douglas C. Saffran
; APPLICANT: Steve Chappell Mitchell
; TITLE OF INVENTION: NOVEL SERPENTINE TRANSMEMBRANE ANTIGENS
; TITLE OF INVENTION: EXPRESSED IN HUMAN CANCERS AND USES THEREOF
; FILE REFERENCE: 129.16USU2
; CURRENT APPLICATION NUMBER: US/09/323,873A
; PRIOR APPLICATION NUMBER: 60/087,520
```

```

; PRIOR FILING DATE: 1998-06-01
; PRIOR APPLICATION NUMBER: 60/091,183
; PRIOR FILING DATE: 1998-06-30
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 339
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-09-323-873A-2
Query Match 30.5%; Score 717; DB 4; Length 339;
Best Local Similarity 54.9%; Pred. No. 8,5e-67;
Matches 130; Conservative 48; Mismatches 59; Indels 0; Gaps 0;

QY 208 LFTLMRGPVVVAISLATFFFLYSFVDVHPVARNQSDPYKPIETVNTKLTPIYATITL 267
DB 67 LFTQWHLPIKIAIILASLTFLLYTLREVIHPILATSHQQYFYKPIPIIVINKVLPMSITLL 126
QY 268 SLVYVLAGLLAAAYOLYYGTRKRRFPPLMTWLCROKGLLSFPFAMVAVASLCUPMR 327
DB 127 ALVYLPGLVNAIVQLNHGTRKRRFPPLMDKMLTRKQGLLSFPFVLAIVSLSPMR 186
QY 328 SERVYFLINMAVOOVANINSENNBEEVRIEMTISFGIMSLGLSLAATVSLPSVSNALN 387
DB 187 SYRYKLLNMAVOOVQNKEDAMIEDVYMBEIVSLGIQVLAITLALATVSLPSVSDSLT 246
QY 388 WRSPFSFIQSTGYVALLISTFHVLLYGMKRAPEEERYAFYTPNPFVLAIVLSIVL 444
DB 247 WRSPFYIQSLGIVSLGLTGLTHALIFAMNMKWDIKQFVYTPPTFMAVFLPIVLI 303

RESULT 4
US-09-685-166A-879
; Sequence 879, Application US/09685166A
; Patent No. 6630305
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Jiang, Yugu
; APPLICANT: Henderson, Robert A.
; APPLICANT: Kaios, Michael D.
; APPLICANT: Fanger, Gary R.
; APPLICANT: Retter, Marc W.
; APPLICANT: Stolk, John A.
; APPLICANT: Day, Craig H.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Carter, Darrick
; APPLICANT: Li, Samuel
; APPLICANT: Wang, Aijun
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Hepler, William
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF PROSTATE CANCER
; FILE REFERENCE: 210121.427C21
; CURRENT APPLICATION NUMBER: US/09/685,166A
; NUMBER OF SEQ ID NOS: 898
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 879
; LENGTH: 339
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-685-166A-879
Query Match 30.5%; Score 717; DB 4; Length 339;
Best Local Similarity 54.9%; Pred. No. 8,5e-67;
Matches 130; Conservative 48; Mismatches 59; Indels 0; Gaps 0;

QY 208 LFTLMRGPVVVAISLATFFFLYSFVDVHPVARNQSDPYKPIETVNTKLTPIYATITL 267
```

Db 67 LFPQMHPIPIAIIASLFLVLTLLREVHPLATSHQYFKYPIPLIVINKVLPMWSTLL 126  
 QY 268 SLVYLGLAAVAQVLYYGNKRRRPRPMLTWTQCKRQGLSFPFPMVAHVASLCPMR 327  
 Db 127 ALVYLPVIAIAIQLNHGKTKKPKFMDLKNMLTKRQFGLSFPFVAHVAIYSLSPMR 186  
 QY 328 SERLPLNMAVQGVHANIEHNSMNEEVMRLIMVYISFGMSGLSLAVTSIPBSNALN 387  
 Db 187 SYRKLLNMAVQGVQONKEDAMTEHVMKKEIYVSLGVALIILALAVTSIPBSDSL 246  
 QY 388 WRPSFIQSLGVALLISFHVLIIGNKRAFESEYRPPNPVNLALVPSVIL 444  
 Db 247 WRREHYIQSLGIVSLIGTIIHLLIANNKMDIDIQFVWYIPPTMLNVAEPLVIL 303

```

RESULT 5
US-09-655-270A-15
Sequence 15, Application US/09655270A
Patent No. 6329151
GENERAL INFORMATION:
APPLICANT: Rouviere, Pierre E.
TITLE OF INVENTION: High Density Sampling of Differentially Expressed Prokaryotic mRNAs
FILE REFERENCE: BC1011 US NA
CURRENT APPLICATION NUMBER: US/09/655,270A
PRIORITY FILING DATE: 2000-09-05
PRIORITY APPLICATION NUMBER: 60/120,702
PRIORITY FILING DATE: 1999-February-19
PRIORITY APPLICATION NUMBER: 60/152,542
PRIORITY FILING DATE: 1999-September-03
NUMBER OF SEQ ID NOS: 37
SOFTWARE: Microsoft Office 97
SEQ ID NO 15
LENGTH: 227
TYPE: PR1
ORGANISM: Rhodococcus erythropolis HL PM-1
US-09-655-270A-15

```

Query Match	6.1%;	Score 144.5;	DB 4;	Length 227;
Best Local Similarity	28.0%;	Prod. No. 6.4e-07;		
Matches	60;	Conservative	33;	Mismatches 88; Indels 33; Gaps 9

  

QY	25	IKDARVTVGVIGSGDPAKSLTIRLCRGYHVIVGSRNPKFASFFPH-----VVD	75
Db	1	MSSSKLAVVG--GTGPGGKGLAYRPAAGWGVIVIGSSASRAEFAALEVRRRAGDGVAYS	58
QY	76	VTIHEDALFKTNIIFFAI-----HREHYSLMDLHLVKGKILI-----DYSNNMRIN	123
Db	59	AAMNBAALADCPILLIVVYDGHRLV--ELAPTRAGKLVVSCNPLGFDGSGAYGID	115
QY	124	QYPSNAEYIASLPDPSILVKGFNVA--WALQPKDASRQVYICSNNIICARQVIE	180
Db	116	VEEGSAAEQRLDIPGATVVAAFHLLSVNLMERE-GP--LPEDVAVCGDGRAGKQEVAR	172
QY	181	LAPQLNFIIP-IDLGSLSARREIENPLALFLTLNR	213
Db	173	LAVAITGRPGIDGALRVARQLEPTAVLVINVR	206

US-09-651-941-17  
 RESULT 6  
 Sequence 17, Application US/09651941  
 Patent No. 6355470  
 GENERAL INFORMATION:  
 APPLICANT: ROUVIER, PIERRE E  
 APPLICANT: WALTERS, DANA M  
 APPLICANT: RAIMER, RUSS  
 TITLE OF INVENTION: Genes Encoding Picti  
 FILE REFERENCE: BC1022 US NA  
 CURRENT APPLICATION NUMBER: US/09/651,941  
 CURRENT FILING DATE: 2000-08-31  
 PRIOR APPLICATION NUMBER: 60/152,545  
 PRIOR FILING DATE: 1999-10-03  
 NUMBER OF SEQ ID NOS: 28

```

: SOFTWARE: Microsoft Office 97
:
: SEQ ID NO 17
:
: LENGTH: 227
:
: TYPE: prt
:
: ORGANISM: Rhodococcus erythropolis HL PM-1
:
: US-03-651-941-17

```

Query Match	6.1%;	Score 14.5;	DB 4.	Length 227;
Best Local Similarity	28.0%;	Pred. No. 6,4e-07;		
Matches 60;	Conservative 33;	Mismatches 88;	Indels 33;	Gaps 9

QY 76 VTHEEDALTKNIIPVAI---HREHYTSLMDREHLVAKILLI-----DVSNMNRIN 123  
Db 59 AADNAGAADCPILLVVPYDGHRLVS---ELAPIFAGKLIVSCNPLGFDKSGAYGLD 115  
QY 124 QYESNAEYTIASLFPDSLIYKGFNVSA---WALQFGKASROYIICSNNTIARQOYTE 180  
Db 116 VEBGSAEOLRDLVPATVAAEFHLSVSNIMEH-GP--LPEDVLVCDDRSKQEVAR 172  
QY 181 LARQLNFIIP-IDGLSSSARBEIENPLLFITLMR 213  
Db 173 LAVAITGRPGIOGALRVARQLEPTAVLINNR 206

```

RESULT 7
US-09-955-597-17
Sequence 17, Application US/0955597
Patent No. 6461856
GENERAL INFORMATION:
APPLICANT: ROUVIER, PIERRE E
APPLICANT: WALTERS, DANA M
APPLICANT: RAINER, RUS
TITLE OF INVENTION: Genes Encoding Picric Acid Degradation
FILE REFERENCE: BC1022 US NA
CURRENT APPLICATION NUMBER: US/09/955,597
CURRENT FILING DATE: 2001-09-17
PRIOR APPLICATION NUMBER: 60/152,545
PRIOR FILING DATE: 1999-10-03
NUMBER OF SEQ ID NOS: 28
SOFTWARE: Microsoft Office 97
SEQ ID NO 17
LENGTH: 227
TYPE: PRT
ORGANISM: Rhodococcus erythropolis HL PM-1
US-09-955-597-17

```

Query March	6.1%	Score 144.5;	DB 4;	Length 227;
Best Local Similarity	28.0%;	Pred. No. 6.4e-07;		
Matches	60;	Conservative	33;	Mismatches 88; Indels 33; Gaps 9
Qy	25	IKDARVTVGVIGSGDFAKSLTIRLIRCGYVVVIGSNPFASEFPFH-----VVD	75	
Db	1	MSSKIAVVG--GTDPQKGLAVRFAAGGVVIGSASARAEALVTRARRGDAVVS	58	
Qy	76	VTHHEALTKNTIIVAI-----HEHTSLMDLPHLLVGLIL-----DYSNNMRIN	123	
Db	59	AADNAAADDCIILIVPYDGRHELVS--ELAPITAGKLVVS CVNPLGFDGSAAGYD	115	
Qy	124	QYFESNAETIASLFPDSILVKGFNVSF--WALQDPCDASKQYVTCSSNIIDARQOYIE	180	
Db	116	VEEGSAABQDRDLVGGATVVAALFHLLSVMNIMEE--GP--LPEDVLVCGDGRSAKQEVAR	172	
Qy	181	LARQLNFIPI-DLQSLSARREIENIPRLFTLMR	213	
Db	173	LAVAITGRGIDGALRVAAQLEPTITAVLVNVR	206	

RESULT 8  
US-08-487-886-2

Sequence 2. Application US/08487886  
Patent No. 5744448  
GENERAL INFORMATION:  
APPLICANT: Kelton, Christie Ann  
APPLICANT: Schweickhardt, Rene Lynn  
APPLICANT: Cheng, Shirley Vui Yen  
APPLICANT: Nugent, No. 5744448een Patrice  
TITLE OF INVENTION: Human Follicle Stimulating  
TITLE OF INVENTION: Hormone Receptor  
NUMBER OF SEQUENCES: 2  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Stephan P. Williams,  
ADDRESSEE: Ares-Serono, Inc.,  
STREET: Exchange Place, 37th floor  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.5" diskette, 1.44 MB, high density  
COMPUTER: IBM PS/2, model 55 SX  
OPERATING SYSTEM: MS-DOS version 4.0  
SOFTWARE: VAX/VMS Massill via Kermit to IBM MS-DOS  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/487,886  
FILING DATE: 07-JUN-1995  
CLASSIFICATION: 514  
PRIORITY APPLICATION DATA:  
APPLICATION NUMBER: 07/670,085  
FILING DATE: 15-MAR-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: Williams, Stephan P.  
REGISTRATION NUMBER: 28546  
REFERENCE/DOCKET NUMBER: US/252  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617) 723-1300  
TELEFAX: (617) 723-8923  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 695  
TYPE: Amino acid  
TOPOLOGY: Linear  
MOLECULE TYPE: protein  
FEATURE:  
NAME/KEY: signal sequence  
LOCATION: -17 to -1  
IDENTIFICATION METHOD: hydrophobic  
FEATURE:  
NAME/KEY: putative amino-terminal extracellular domain  
LOCATION: 1 to 349  
IDENTIFICATION METHOD: similarity with other  
IDENTIFICATION METHOD: dimeric glycoprotein receptor extracellular  
IDENTIFICATION METHOD: domains, hydrophilic  
FEATURE:  
NAME/KEY: transmembrane domain  
LOCATION: 350 to 613  
IDENTIFICATION METHOD: similarity to other G  
IDENTIFICATION METHOD: protein-coupled receptor transmembrane domains  
FEATURE:  
NAME/KEY: putative transmembrane region I  
LOCATION: 350 to 370  
IDENTIFICATION METHOD: similarity to other G  
IDENTIFICATION METHOD: protein-coupled receptor transmembrane regions,  
IDENTIFICATION METHOD: hydrophobic, about 20-23 amino acids in length  
FEATURE:  
NAME/KEY: putative transmembrane region II  
LOCATION: 382 to 404  
IDENTIFICATION METHOD: similarity to other G  
IDENTIFICATION METHOD: protein-coupled receptor transmembrane regions,  
IDENTIFICATION METHOD: hydrophobic, about 20-23 amino acids in length  
FEATURE:  
NAME/KEY: putative transmembrane region III  
LOCATION: 427 to 448

IDENTIFICATION METHOD: similarity to other G  
IDENTIFICATION METHOD: protein-coupled receptor transmembrane regions,  
IDENTIFICATION METHOD: hydrophobic, about 20-23 amino acids in length  
FEATURE:  
NAME/KEY: putative transmembrane region IV  
LOCATION: 469 to 491  
IDENTIFICATION METHOD: similarity to other G  
IDENTIFICATION METHOD: protein-coupled receptor transmembrane regions,  
IDENTIFICATION METHOD: hydrophobic, about 20-23 amino acids in length  
FEATURE:  
NAME/KEY: putative transmembrane region V  
LOCATION: 512 to 533  
IDENTIFICATION METHOD: similarity to other G  
IDENTIFICATION METHOD: protein-coupled receptor transmembrane regions,  
IDENTIFICATION METHOD: hydrophobic, about 20-23 amino acids in length  
FEATURE:  
NAME/KEY: putative transmembrane region VI  
LOCATION: 557 to 580  
IDENTIFICATION METHOD: similarity to other G  
IDENTIFICATION METHOD: protein-coupled receptor transmembrane regions,  
IDENTIFICATION METHOD: hydrophobic, about 20-23 amino acids in length  
FEATURE:  
NAME/KEY: putative transmembrane region VII  
LOCATION: 592 to 613  
IDENTIFICATION METHOD: similarity to other G  
IDENTIFICATION METHOD: protein-coupled receptor transmembrane regions,  
IDENTIFICATION METHOD: hydrophobic, about 20-23 amino acids in length  
FEATURE:  
NAME/KEY: putative carboxy-terminal intracellular  
LOCATION: 614 to 678  
US-08-487-886-2

Query Match 4.6%; Score 107.5; DB 1; Length 695;  
Best Local Similarity 17.0%; Pred. No. 0.028;  
Matches 84; Conservative 77; Mismatches 148; Indels 185; Gaps 17;

QY 79 HEDALTKNIIFVAIHEHYTSIMDLRLVVG-----KILIDVSNMR 121  
DB 98 HEIRIEKANNL-LYINPEAFQNLPNLYLLISNTGKHLPDVKHLSLQKVLIDIQDNIN 156  
QY 122 INQYPSNABYLASLEPDSLLYKGFNVYSAMQLQCPKASROYICSNNGARQVIEL 181  
DB 157 IH-----TIKNSFVGLSFESEVILML--NKGIOEINCA----- 189  
QY 182 AROLNPIPIDLGSLSSAREIENLPLFTLMRGPVVAISLATPFPLYSFVADVIHPYAR 241  
DB 190 ---FNGTQDELNLSDNNLLELPNDVFHGASGPVILDISRTIRSLPSYGLNKKRA 246  
QY 242 NQOSDYKPIPIELVNTL-PIYATLLSLVY----- 271  
DB 247 RSTYNNKLP-----TLKRLVALMEASLYPSHCCAFANMRQISELHPICNKSILROE 300  
QY 272 -----LAGLAAAYOLVYGTCKRRF-----PPLFTWLQCRKOL 305  
DB 301 VDMYQTRGQBSLADNESSYSRGEDMTYTFEDDLCEVVDVYNCSPEPDANFCEDIM 360  
QY 306 G---LISFPAMVA-----YSLCPMERSRYFLNNAYQCVHANTN 347  
DB 361 GYNILRLVLIWFSILATIGNIIVLVITTSQYKLVF-----RFLMCLNLFAD----- 408  
QY 348 SWESEVWMIEMVYISGIMSLGLSLAVTSIPVSNALNMR-----FSFIQSTL 398  
DB 409 -----LCIGIVLLIASVDITKQYANVAIDWGTGACGDAAGFFIYASRL 455  
QY 399 G---YVALLISFTH-----VLIYGRKRAFEERYRFPYPPNFV-IA 435  
DB 456 SVYTLAIATLERWHTTHAMQLDCKYQLRRASVVMGWIFAPAAALPFIIGLISYMKYS 515  
QY 436 LVLPSTVILDLQ 449  
DB 516 ICLPNDIDSPLSQL 529

```

1      RESULT 9
2      US-08/482-855-2
3      Sequence 2, Application US/08482855
4      Patent No. 6121016
5      GENERAL INFORMATION:
6      APPLICANT: Kelton, Christie Ann
7      APPLICANT: Schwelckhardt, Rene Lynn
8      APPLICANT: Cheng, Shirley Vii Yen
9      APPLICANT: Nugent, No. 6121016en Patrice
10     TITLE OF INVENTION: Human Follicle Stimulating
11     TITLE OF INVENTION: Hormone Receptor
12     NUMBER OF SEQUENCES: 2
13     CORRESPONDENCE ADDRESS:
14     ADDRESSEE: Stephan P. Williams,
15     ADDRESSEE: Ares-Serono, Inc.
16     STREET: Exchange Place, 37th floor
17     CITY: Boston
18     STATE: MA
19     COUNTRY: USA
20     ZIP: 02109
21     COMPUTER READABLE FORM:
22     MEDIUM TYPE: 3.5" diskette, 1.44 MB, high density
23     COMPUTER: IBM ps/2, model 55 SX
24     OPERATING SYSTEM: MS-DOS version 4.0
25     SOFTWARE: VAX/VMS Massail via Kermit to IBM MS-DOS
26     CURRENT APPLICATION DATA:
27     APPLICATION NUMBER: US/08/482,855
28     FILING DATE: 07-JUN-1995
29     CLASSIFICATION: 435
30     PRIOR APPLICATION DATA:
31     APPLICATION NUMBER: 07/670,085
32     FILING DATE: 15-MAR-1991
33     ATTORNEY/AGENT INFORMATION:
34     NAME: Williams, Stephan P.
35     REGISTRATION NUMBER: 28546
36     REFERENCE/DOCKET NUMBER: US/252
37     TELECOMMUNICATION INFORMATION:
38     TELEPHONE: (617) 723-1300
39     TELEFAX: (617) 723-8923
40     INFORMATION FOR SRO ID NO: 2:
41     SEQUENCE CHARACTERISTICS:
42     LENGTH: 695
43     TYPE: Amino acid
44     TOPOLOGY: Linear
45     MOLECULE TYPE: protein
46     FEATURE:
47     NAME/KEY: signal sequence
48     LOCATION: -17 to -1
49     IDENTIFICATION METHOD: hydrophobic
50     FEATURE:
51     NAME/KEY: putative amino-terminal extracellular domain
52     LOCATION: 1 to 349
53     IDENTIFICATION METHOD: similarity with other
54     IDENTIFICATION METHOD: dimeric glycoprotein receptor extracellular
55     IDENTIFICATION METHOD: domains, hydrophilic
56     FEATURE:
57     NAME/KEY: transmembrane domain
58     LOCATION: 350 to 613
59     IDENTIFICATION METHOD: similarity to other G
60     IDENTIFICATION METHOD: protein-coupled receptor transmembrane domains
61     FEATURE:
62     NAME/KEY: putative transmembrane region I
63     LOCATION: 350 to 370
64     IDENTIFICATION METHOD: similarity to other G
65     IDENTIFICATION METHOD: protein-coupled receptor transmembrane regions
66     IDENTIFICATION METHOD: hydrophobic, about 20-23 amino acids in length
67     FEATURE:
68     NAME/KEY: putative transmembrane region II
69     LOCATION: 382 to 404
70     IDENTIFICATION METHOD: similarity to other G
71     IDENTIFICATION METHOD: protein-coupled receptor transmembrane regions
72     IDENTIFICATION METHOD: hydrophobic, about 20-23 amino acids in length

```

```

1  FEATURE:
2  NAME/KEY: putative transmembrane region III
3  LOCATION: 427 to 448
4  IDENTIFICATION METHOD: similarity to other G
5  IDENTIFICATION METHOD: protein-coupled receptor transmembrane regions,
6  IDENTIFICATION METHOD: hydrophobic, about 20-23 amino acids in length
7  FEATURE:
8  NAME/KEY: putative transmembrane region IV
9  LOCATION: 469 to 491
10 IDENTIFICATION METHOD: similarity to other G
11 IDENTIFICATION METHOD: protein-coupled receptor transmembrane regions,
12 IDENTIFICATION METHOD: hydrophobic, about 20-23 amino acids in length
13 FEATURE:
14 NAME/KEY: putative transmembrane region V
15 LOCATION: 512 to 533
16 IDENTIFICATION METHOD: similarity to other G
17 IDENTIFICATION METHOD: protein-coupled receptor transmembrane regions,
18 IDENTIFICATION METHOD: hydrophobic, about 20-23 amino acids in length
19 FEATURE:
20 NAME/KEY: putative transmembrane region VI
21 LOCATION: 557 to 580
22 IDENTIFICATION METHOD: similarity to other G
23 IDENTIFICATION METHOD: protein-coupled receptor transmembrane regions,
24 IDENTIFICATION METHOD: hydrophobic, about 20-23 amino acids in length
25 FEATURE:
26 NAME/KEY: putative transmembrane region VII
27 LOCATION: 592 to 613
28 IDENTIFICATION METHOD: similarity to other G
29 IDENTIFICATION METHOD: protein-coupled receptor transmembrane regions,
30 IDENTIFICATION METHOD: hydrophobic, about 20-23 amino acids in length
31 FEATURE:
32 NAME/KEY: putative carboxy-terminal intracellular
33 NAME/KEY: domain
34 LOCATION: 614 to 678
35 US-08-482-855-2
36
37 Query Match 4.6%; Score 107.5; DB 3; Length 695;
38 Best Local Similarity 17.0%; Pred. No. 0.028;
39 Matches 84; Conservative 77; Mismatches 148; Indels 185; Gaps 17
40
41 QY 79 HEDALTKTIIIVAIHREHYTSLMDRLSLVG-----KILIDVENNR 121
42 DB 98 HEIRIKANNLTVINPEAFONLPTQLILSNIGHLDPVKHISLQKVLIDIDNIN 156
43
44 QY 122 INQYPSNAEYIASLPFDSLIYKGNVNSAMALQPKDASROYIYCSNNIQAQOVIETL 181
45 DB 157 IH-----TIERSFVGLSFESEYIWL-----NKNGIOEIHNCA----- 189
46
47 QY 182 AROQNIPIFDLSLSAREINLPRLPTLMRGVVAVALAFPFYSPVRDVIHVAR 241
48 DB 190 ---FNGTQDELNLSDNNLELPDVPFHGASGVYIIDIRTHISLPSIGLENLKRA 246
49
50 QY 242 NQSDPYKPIEIVNKLTPVAITLSLVY----- 271
51 DB 247 RSTYNNKLP-----TLEKVALMEASLITVPSHCACAFANWRQISELHPICNKSILROE 300
52
53 QY 272 -----LAGLAAAYQLYGKTKRRF-----PMTLETLQCKRL 305
54 DB 301 VDIYMTQGRGSLAEDNESSYSGFDWTYTFEYDLCNEVDVTCSPKPDANPCEDIM 360
55
56 QY 306 G-----LISFFAMVHA-----YSLCPMRSEERYLFLANAAQOVHANEN 347
57 DB 361 GNNILKVLWFLISLAINIGNIIIVILITTSQYKLTVP-----RFLMGNLAFND----- 408
58
59 QY 348 SNNEEVARIEMYSIFGMSIGLSLAVTSIPSVSNAALNRE-----FSFIQSTL 398
60 DB 409 -----LCIGIYLLIASVDIHKTSQYHNALIDMQGACDAGFEFTYFASEL 455
61
62 QY 399 G---YVALLISTFH-----VLIIYGNKAFEEZYRFFYTPPV-IA 435
63 DB 456 SYVTLTATLEKHHITTHAMQDCKYQULRHASAMWGMIFAPAAALLPFIIGISSTMYKS 515
64
65 QY 436 LVLPSTVILDLQL 449

```

Db 516 ICLPMDIDPSLQ 529

RESULT 10

US-08-474-986-2

GENERAL INFORMATION:

APPLICANT: Kelton, Christie Ann

Schweickhardt, Rene Lynn

Cheng, Shirley Wei Yen

Nugent, No. 637271leen Patrice

TITLE OF INVENTION: Human Follicle Stimulating

Hormone Receptor

NUMBER OF SEQUENCES: 2

CORRESPONDENCE ADDRESS:

ADDRESSEE: Stephan P. Williams,

Ares-Serono, Inc.

STREET: Exchange Place, 37th floor

CITY: Boston

STATE: MA

COUNTRY: USA

ZIP: 02109

COMPUTER READABLE FORM:

MEDIUM TYPE: 3.5" diskette, 1.44 MB, high density

COMPUTER: IBM PS/2, model 55 SX

OPERATING SYSTEM: MS-DOS version 4.0

SOFTWARE: VAX/VMS Mass11 via Kermit to IBM MS-DOS

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/474,986

FILING DATE: 07-Jun-1995

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 07/670,085

FILING DATE: <Unknown>

ATTORNEY/AGENT INFORMATION:

NAME: Williams, Stephan P

REGISTRATION NUMBER: 28546

REFERENCE/DOCKET NUMBER: US/252

TELECOMMUNICATION INFORMATION:

TELEPHONE: (617) 723-1300

TELEFAX: (617) 723-8923

LOCATION: 614 to 678

SEQUENCE DESCRIPTION: SEQ ID NO: 2:

US-08-474-986-2

Query Match 4.6%; Score 107.5; DB 4; Length 695;  
Best Local Similarity 17.0%; Pred. No. 0.028;  
Matches 84; Conservative 77; Mismatches 148; Indels 185; Gaps 17;

Db 79 HEDALTKNTIIFAIHREHTSLMDLRHLVG-----KIIDVSNMR 121  
98 HETIRKANNL-LVINPEAFONLPNLYLLISNTGIKHLDVHKHSLOKVLDDQDNIN 156  
122 INQYPSNAEYLAIPDPSLIYKGFNVVSAWALQSPKASROVYICSNINQARQVTEL 181  
157 IH-----TIRNSFVGISFESVILML-----NKNGIGIEINHC----- 189

Db 182 AROLATIPDLGSLSSAREIENIPLRLFTLMKGPVVAASLATFFLYSFVADVHPAR 241  
190 ---FNGTDLDELMLSDNNMLELPNDVFAGSGPVILDISRIRHSLPSYGENIKQRA 246  
242 NQSDFKPIEIVNKTLL-PIVAITLLSVY----- 271  
247 RSTYNNKCLP-----TLKCVALLMASILTPSHCCAFANWRQISELHPICKSLROE 300

Db 272 -----LAGLMAAYQLYYGTKRRF-----PPLLETWLQCRKOL 305  
301 VDYWTQTRGQRSSIAEDNESSYSGRPMYTERFDYDLCEVVDVTCSPKPDAFNCEEDIM 360  
306 G-----LISFFAMVVA-----YSLCLPMRSEYYLFINMAYQOVANINEN 347

Db 361 GYMLIKVLIWFISILATIGNITIVVILITTSQYKLTVP-----RFLMCLALFAD----- 408

Qy 348 SWNEEYWRLEWYISFGINSGLSLAATSPISVSNALNMR-----PSFIQSTL 398  
Db 409 -----LCIGIYLLIASVDIHTKSGYHVAIDMQGACDADAAGFTVPASL 455

Qy 399 G---YVALLISTF-----VLIYKRPAREEYRYPPNIV-IA 435  
Db 456 SVYTLTAITLERWHTTHAMQDLCKYQLRHASVWGMIFAPAAALPFIIGISYWKVS 515

Qy 436 LVPSTVLDLQ 449  
Db 516 ICLPMDIDPSLQ 529

RESULT 11

US-09-328-873A-20

Sequence 20; Application US/0932873A

Patent No. 6329503

GENERAL INFORMATION:

APPLICANT: Daniel E. Afar

APPLICANT: Rene S. Hubert

APPLICANT: Kahan Leong

APPLICANT: Arthur B. Raitano

APPLICANT: Douglas C. Saffran

APPLICANT: Steve Chappell Mitchell

TITLE OF INVENTION: NOVEL SERPENTINE TRANSMEMBRANE ANTIGENS

FILE REFERENCE: 129.16USU2

CURRENT APPLICATION NUMBER: US/09/323,873A

CURRENT FILING DATE: 1999-06-01

PRIOR APPLICATION NUMBER: 60/087,520

PRIOR FILING DATE: 1998-06-01

PRIOR APPLICATION NUMBER: 60/091,183

PRIOR FILING DATE: 1998-06-30

NUMBER OF SEQ ID NOS: 32

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 20

LENGTH: 34

TYPE: PRT

ORGANISM: Artificial Sequence

FEATURE:

OTHER INFORMATION: STEAP-1 PEPTIDE

US-09-323-873A-20

Query Match 4.6%; Score 107; DB 4; Length 34;  
Best Local Similarity 58.8%; Pred. No. 0.0003; 9; Indels 0; Gaps 0;  
Matches 20; Conservative 5; Mismatches 9;

Qy 326 RRSERYLFLNMAYQOVANINENSWNEEYWRIM 359  
Db 1 RRSYRYLNLWVAYQOVQNKEDAMIEHDVWRMEI 34

RESULT 12

US-09-328-352-8165

Sequence 8165; Application US/09328352

Patent No. 6562958

GENERAL INFORMATION:

APPLICANT: Gary L. Breton et al.

TITLE OF INVENTION: NOCATEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER

FILE REFERENCE: GTC09-03PA

CURRENT APPLICATION NUMBER: US/09/328,352

CURRENT FILING DATE: 1999-06-04

NUMBER OF SEQ ID NOS: 8252

SEQ ID NO 8165

LENGTH: 940

TYPE: PRT

ORGANISM: Acinetobacter baumannii

US-09-328-352-8165

Query Match 4.4%; Score 102.5; DB 4; Length 940;  
Best Local Similarity 19.5%; Pred. No. 0.15; Indels 73; Gaps 17;  
Matches 71; Conservative 74; Mismatches 146;



QY 88 IIFVAIHREHYTSIMDLRHLLVSKILLIDVSNMNRINOYPESNAEYLAFL-----PDSLI 142  
DB 262 VVFLTLVSHHGV-----ITSLTASILFTLLATVFLSKQOATVLAIALGMAVADLV 319  
QY 143 VKGF--NVVSAMALQUGPKDASROVYICSNNIQARQVIELARQIN-----FIPIDLSLS 196  
DB 320 IQYAPDVVFLFS-----YLVNLAVALVNFILQPKILNQAFALMIGSSA 368  
QY 197 SAREIENLPLRLFTL--MRGPVVAISLATEFFLYSFVDVHPFARNQOSPFIKPIEIV 255  
DB 369 IAFVAE--PAKEDTIDM-----ILMHLIFLWLSRVYSONISRVSEHKOGLP----- 418  
QY 256 NKTLPVATLTLISLVLAGLAAAYQLYYGRK-----YRRPPLMETWLOCRK 303  
DB 419 ----FLDVGILFNVPLVIGFTLHAYLVHESIOALTGAVALAGYAVUTWIK--KTHP 471  
QY 304 QUGLISFFPAMVHA--YSLCLPMRSEERYLFLNNAVQOVHANIEVNEEWEIEMVYS 362  
DB 472 QLSVLAKSFFILAVAFPLIFPLAKGAWTAIGVAOCTALIV--MGVTERYLSRYL- 527  
QY 363 FQIMSLGSLLAATVSISSVSNALNW--REPSFIQSTIGYVALLISTHVLITGKKAPE 420  
DB 528 -GVILVLSLALFYQV-----WANEFPPTLSTISYIAIQAFTSAFYLLQVNSK--E 575  
QY 421 EBYV 424  
DB 576 QRYF 579

## RESULT 13

US-09-316-083-3  
; Sequence 3, Application US/09316083A  
; Patent No. 6280942  
; GENERAL INFORMATION:  
; APPLICANT: The Institute of Physical and Chemical Research  
; TITLE OF INVENTION: Endonuclease  
; FILE REFERENCE: PH-651  
; CURRENT APPLICATION NUMBER: US/09/316, 083A  
; CURRENT FILING DATE: 1999-05-20  
; EARLIER APPLICATION NUMBER: JP98/141861  
; EARLIER FILING DATE: 1998-05-22  
; NUMBER OF SEQ ID NOS: 38  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 3  
; LENGTH: 476  
; TYPE: PRT  
; ORGANISM: *Saccharomyces cerevisiae*  
US-09-316-083-3

Query Match 4.3%; Score 101; DB 3; Length 476;  
Best Local Similarity 20.7%; Pred. No. 0.075;  
Matches 52; Conservative 48; Mismatches 85; Indels 66; Gaps 12;

QY 62 NPKASEFPFHVVDVTHHEDALTKNI-----IFVAIHREHYT-----SLMD 103  
DB 254 NPYFVNAFSINI-----KTNLAKKIFNTYKLSYDYKINQINNHIPYNYLK 302  
QY 104 LRHLVGLKILLIDVSNMNRINOYPESNAEYLAFL--PDSLIYKGFNVVSAMALQUGPKDA 161  
DB 303 INNKLPKINMIDIKNNVYLAGFTADGSLSMYNNPKOTLLFKNM----- 347  
QY 162 SROYVICSNNIQARQVIELARQINFIPIIDGSLSSAREIENLPLRLFTLMRGPVVAIS 221  
DB 348 -RPSVVIS--QVETREKELIYLIOE-----SFDL--SISNVKVGGRKJDFLFRITDELK 400  
QY 222 LATFPFLYSFVDVHPFARNQOSDPYKIP-----IEIVNKTLPVATLTLISLVLAGL 277  
DB 401 ----FIYF--DKFLPHDNKQFNYIKFRFTFIKSYNNMNRVGLVLS--YINNIKI 451  
QY 278 AAYQLYYGRY 288  
DB 452 DNYDYYYNRY 462

## RESULT 14

US-09-933-700-3  
; Sequence 3, Application US/09933700  
; Patent No. 6528296  
; GENERAL INFORMATION:  
; APPLICANT: The Institute of Physical and Chemical Research  
; TITLE OF INVENTION: Endonuclease  
; FILE REFERENCE: PH-651  
; CURRENT APPLICATION NUMBER: US/09/933, 700  
; CURRENT FILING DATE: 2001-08-20  
; PRIOR APPLICATION NUMBER: 09/316, 083  
; PRIOR FILING DATE: 1999-05-20  
; NUMBER OF SEQ ID NOS: 38  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 3  
; LENGTH: 476  
; TYPE: PRT  
; ORGANISM: *Saccharomyces cerevisiae*  
US-09-933-700-3

Query Match 4.3%; Score 101; DB 4; Length 476;  
Best Local Similarity 20.7%; Pred. No. 0.075;  
Matches 52; Conservative 48; Mismatches 85; Indels 66; Gaps 12;

QY 62 NPKASEFPFHVVDVTHHEDALTKNI-----IFVAIHREHYT-----SLMD 103  
DB 254 NPYFVNAFSINI-----KTNLAKKIFNTYKLSYDYKINQINNHIPYNYLK 302  
QY 104 LRHLVGLKILLIDVSNMNRINOYPESNAEYLAFL--PDSLIYKGFNVVSAMALQUGPKDA 161  
DB 303 INNKLPKINMIDIKNNVYLAGFTADGSLSMYNNPKOTLLFKNM----- 347  
QY 162 SROYVICSNNIQARQVIELARQINFIPIIDGSLSSAREIENLPLRLFTLMRGPVVAIS 221  
DB 348 -RPSVVIS--QVETREKELIYLIOE-----SFDL--SISNVKVGGRKJDFLFRITDELK 400  
QY 222 LATFPFLYSFVDVHPFARNQOSDPYKIP-----IEIVNKTLPVATLTLISLVLAGL 277  
DB 401 ----FIYF--DKFLPHDNKQFNYIKFRFTFIKSYNNMNRVGLVLS--YINNIKI 451  
QY 278 AAYQLYYGRY 288  
DB 452 DNYDYYYNRY 462

## RESULT 15

US-09-252-991A-21494  
; Sequence 21494, Application US/09252991A  
; Patent No. 6551795  
; GENERAL INFORMATION:  
; APPLICANT: Marc J. Rubenfield et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
; FILE REFERENCE: 107196.136  
; CURRENT APPLICATION NUMBER: US/09/252, 991A  
; CURRENT FILING DATE: 1999-02-18  
; PRIOR APPLICATION NUMBER: US 60/074, 788  
; PRIOR FILING DATE: 1998-02-18  
; PRIOR APPLICATION NUMBER: US 60/094, 190  
; PRIOR FILING DATE: 1998-07-27  
; NUMBER OF SEQ ID NOS: 33142  
; SEQ ID NO 21494  
; TYPE: PRT  
; ORGANISM: *Pseudomonas aeruginosa*  
US-09-252-991A-21494

Query Match 4.3%; Score 101; DB 4; Length 724;  
Best Local Similarity 20.5%; Pred. No. 0.14;  
Matches 72; Conservative 57; Mismatches 105; Indels 118; Gaps 16;

```
QY 132 YLASLPDSLIYKGFVNSAMALQGFKQASRQVYICSNNTQA-----RQVIELARQLN 186
Db 298 YLSVLPDLSL-SKTLAMVMAVYLVCGTLPSA-----LCVISLSLSGPHRQRALDILRROA 352
QY 187 FIPIDL-GSLSAREIENLPLRLFTLMRGPVVAISLA-----TFPELYSFVRDVIHP 238
Db 353 FRPLWLIGSLALGEVAHDPRLIAGLGEHTSICSLTANASALFTALFVWRFRRIAH- 411
QY 239 YARNQ-----SDPYKIPi-----EIVNK 257
Db 412 LIRNOPLERLRKRSLHDLVQVLGSLMFVPLVVLVGISLPATFVSAGDSSSALRRALVCA 471
QY 258 TLPIVAITLLSLVYLAGLLAAAYQLYGTRRFPPEMLETWLQCRKQGLLSFPFAMVHV 317
Db 472 VLAVVAMTVIGLI-----RRSSRVGAGPRRSAPYIQ-----LQSFGYTLIHI 515
QY 318 AVSLCLPMRRSERVYFLNNAVYQQVHANINENSWNEEFWRLEM-YISFGIM----- 366
Db 516 FVVL-----FIEVALRWGMSLIR-YAEGEGQISMKVVSFGTTLVAMLIWIL 564
QY 367 -----SLGL-----LSLLAVTSIPVSNALNWRREFSFIQSTLGVALIIS 406
Db 565 TDTAIQHSJLGGKSRPNTALTMPLIRNVL-----FATIAVIALIVA 608
```

Search completed: March 1, 2004, 23:40:58  
Job time : 48 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 1, 2004, 19:51:19 ; Search time 40 Seconds

(without alignments)  
590,996 Million cell updates/sec

Title: US-09-455-486-6

Perfect score: 2351

Sequence: 1 MESISMWGSPKSLSTCLPN.....ALVLPSTVILDLQLCKYPD 454

Scoring table: BLASTUM62

Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_42.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	717	30.5	339	1	STEA_HUMAN
2	133	5.7	223	1	FARE_MERLA
3	120.5	5.1	232	1	FARE_MERTH
4	111.5	4.7	695	1	FSHR_MACPA
5	109.5	4.6	206	1	Y538_PASNU
6	109	4.6	693	1	FSHR_CHICK
7	109	4.6	712	1	FREE6_YEAST
8	107.5	4.6	629	1	FREE7_YEAST
9	107.5	4.6	695	1	FSHR_HUMAN
10	104.5	4.4	695	1	FSHR_PIG
11	104	4.4	396	1	SOTB_ECO57
12	104	4.4	396	1	SOTB_ECOLI
13	103	4.4	465	1	NPTI_MOUSE
14	102.5	4.4	391	1	NCCC_NEPOL
15	102	4.3	220	1	Y304_BRUME
16	102	4.3	220	1	Y301_BRUST
17	100	4.2	476	1	RS3_SACBA
18	99.5	4.2	452	1	NU4M_BACBA
19	99	4.2	365	1	SPRI_HUMAN
20	98	4.2	3411	1	POLG_YEYV1
21	98	4.2	3411	1	POLG_YEYF2
22	97.5	4.1	268	1	YD49_AOUAE
23	97.5	4.1	695	1	FSHR_BOVIN
24	97	4.1	218	1	FSHR_BOVIN
25	97	4.1	692	1	FSHR_XANNC
26	96.5	4.1	686	1	FSHR_RAT
27	96	4.1	261	1	PRE1_YEAST
28	95	4.0	202	1	PHSC_ECOLI
29	95	4.0	393	1	NUCC_ARATH
30	95	4.0	393	1	NUCC_SPIOL
31	95	4.0	692	1	FSHR_MOUSE
32	94.5	4.0	492	1	SECY_CYAPA
33	94.5	4.0	695	1	FSHR_SHEEP

34	93.5	4.0	365	1	SPRI_MOUSE	Q8bfc3	mus musculus
35	93.5	4.0	452	1	NU4M_BACFL	O47423	branchiost
36	93.5	4.0	499	1	FSHR_PLAY7	O15770	plasmidum
37	93.5	4.0	530	1	AAAI_MOUSE	Q914h8	mus musculus
38	93.5	4.0	1233	1	VIL1_REOVD	P15024	reovirus (t
39	93	4.0	388	1	HMC3_DESVH	P33390	desulfovibr
40	93	4.0	459	1	NU4M_BOVIN	P03910	bos taurus
41	93	4.0	521	1	YT25_CABEL	Q10934	caenorhabdi
42	92.5	3.9	490	1	GTRI_CHICK	P46896	gallus gall
43	92.5	3.9	500	1	NU4C_OENHO	P58419	oenothera h
44	92.5	3.9	883	1	YHL6_YEAST	P38781	saccharomyc
45	92.5	3.9	1769	1	YXK9_YEAST	P42945	saccharomyc

## ALIGNMENTS

RESULT 1	ID	STEA_HUMAN	STANDARD	PRT	339 AA.
AC	09H8A	O95034			
DT	16-OCT-2001	(Rel. 40, Created)			
DT	16-OCT-2001	(Rel. 40, Last sequence update)			
DT	15-MAR-2004	(Rel. 43, Last annotation update)			
DE	Six transmembrane epithelial antigen of prostate.				
GN	STEAP OR STEAP1.				
OS	Homo sapiens (Human).				
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;				
OC	Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.				
OX	NCBI_TaxID=9606;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RX	MEDLINE=20056277; PubMed=10588738;				
RA	Hubert R.S., Vivanco I., Chen E., Rastegar S., Leong K.,				
RA	Mitchell S.C., Madraswala R., Zhou Y., Kuo J., Raitano A.B.,				
RA	Uakobovits A., Salfran D.C., Afar D.E.H.;				
RT	"STEAP: a prostate-specific cell-surface antigen highly expressed in				
RT	human prostate tumors.";				
RT	Proc. Natl. Acad. Sci. U.S.A. 96:14523-14528(1999).				
RN	[2]				
RP	SEQUENCE FROM N.A.				
RA	Abu-Threideh J., Stoneking T., Langston Y., Maupin R.;				
RA	Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.				
RN	[3]				
RP	SEQUENCE FROM N.A.				
RC	TISSUE=Skin;				
RX	MEDLINE=22388257; PubMed=12477932;				
RA	Straussberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,				
RA	Klaunig R.D., Collins F.S., Wagner L., Sherman C.M., Schuler G.D.,				
RA	Altschul S.F., Zeeberg B., Buettow K.H., Schaefer C.F., Bat N.K.,				
RA	Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.;				
RA	Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,				
RA	Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,				
RA	Brownstein M.J., Usdin T.B., Toshitoki S., Carinoni P., Prange C.,				
RA	Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,				
RA	Bosak S.A., Keshan P.J., McKernan K.O., Malek J.A., Gunaratne P.H.,				
RA	Richard S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulix S.W.,				
RA	Villalón D.K., Mizny D.M., Sodergren E.J., Lu X., Gibbs R.A.,				
RA	Fahy J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,				
RA	Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,				
RA	Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,				
RA	Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,				
RA	Butterfield Y.S.N., Krzyzinski W.T., Skalska U., Smalins D.E.,				
RA	Schnerch A., Schein J.E., Jones S.J.W., Marra M.A.;				
RT	"Generation and initial analysis of more than 15,000 full-length				
RT	human and mouse cDNA sequences.";				
RT	Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).				
CC	-1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).				
CC	-1- TISSUE SPECIFICITY: Highly expressed in prostate tumors.				
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration				
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -				
CC	the European Bioinformatics Institute. There are no restrictions on its				

CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL: AF186249; AAF17479.1; -  
 DR EMBL: AC005053; AAC79150.1; ALT\_INIT.  
 DR EMBL: AC004969; AAD15620.2; -  
 DR EMBL: BC011802; AAH1802.1; -  
 DR Gene; HGNC:11378; STEAP.  
 DR MIM: 604415; -  
 DR GO: GO:0005887; C: integral to plasma membrane; TAS.  
 DR GO: GO:0005911; C: intercellular junction; TAS.  
 DR GO: GO:0015267; F: channel/pore class transporter activity; TAS.  
 KW Transmembrane, Antigen.  
 FT TRANSMEM 71 91 POTENTIAL.  
 FT TRANSMEM 119 139 POTENTIAL.  
 FT TRANSMEM 164 184 POTENTIAL.  
 FT TRANSMEM 218 238 POTENTIAL.  
 FT TRANSMEM 258 278 POTENTIAL.  
 FT TRANSMEM 291 311 POTENTIAL.  
 SQ SEQUENCE 339 AA; 39851 MW; 55443A170C870387 CRC64;

Query Match 30.5%; Score 717; DB 1; Length 339;  
 Best Local Similarity 54.9%; Pred. No. 66-48;  
 Matches 130; Conservative 48; Mismatches 59; Indels 0; Gaps 0;

QY 208 LFTLRGPRVVAISLSTFFELYSFVDVHPAPNQGDFKIPRIYNTKLPIVATILL 267  
 DB 67 LPPQWMLPTKIAIAISLSTFFELYSFVDVHPAPNQGDFKIPRIYNTKLPIVATILL 126  
 QY 268 SLVYLAGLAAVYQVYGTGKRRFPWLETLQCRKQGLSFFPAPVAVASLCPMR 327  
 DB 127 ALVYLPVIAVQLHNGTKYKKFPHMLDKMLTKKQGLSFFPAPVAVASLCPMR 186  
 QY 328 SRRYFLNAAVYQVYGTGKRRFPWLETLQCRKQGLSFFPAPVAVASLCPMR 387  
 DB 187 SRRYFLNAAVYQVYGTGKRRFPWLETLQCRKQGLSFFPAPVAVASLCPMR 246  
 QY 388 KREFSFIQSTLGYVALLSTFFVLYGWKRAFEERYFPFPPVLAIVPSIYL 444  
 DB 247 WEFHFIQSTLGYVALLSTFFVLYGWKRAFEERYFPFPPVLAIVPSIYL 303

RESULT 2  
 FARE METUA STANDARD; PRT; 223 AA.  
 ID FARE METUA STANDARD; PRT; 223 AA.  
 AC 058896;  
 DT 01-NOV-1997 (Rel. 35, Last Created)  
 DT 10-OCT-2003 (Rel. 42, Last annotation update)  
 DE Putative F420-dependent NADP reductase (EC 1.-.-.-).  
 GN MJ1501.  
 OS Methanococcus jannaschii.  
 OC Archaea; Euryarchaeota; Methanococci; Methanococcales;  
 OC Methanocaldococcaceae; Methanocaldococcus.  
 OX NCBI\_TaxId=2190;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=JAL-1 / DSM 2661 / ATCC 43067;  
 RX MEDLINE=96337999; PubMed=8688087;  
 RA Bult C.J., White O., Olsen G.J., Zhou L., Fleischmann R.D.,  
 RA Sutton G.G., Blake J.A., Fitzgerald L.M., Clayton R.A., Gocayne J.D.,  
 RA Kerlavage A.R., Dougherty B.A., Tomb J.-F., Adams M.D., Reich C.I.,  
 RA Overbeek R., Kirkness E.F., Weinstock K.G., Merrick J.M., Glodek A.,  
 RA Scott J.L., Geoghegan N.S.M., Weidman J.F., Fultman J.L., Nguyen D.,  
 RA Utterback T.R., Kelley J.M., Peterson J.D., Sadow P.W., Hanna M.C.,  
 RA Cotton M.D., Roberts K.M., Hurst M.A., Kaine B.P., Borodovsky M.,  
 RA Klenk H.-P., Raaser C.M., Smith H.O., Koese C.R., Venter J.C.,  
 RT "Complete genome sequence of the methanogenic archaeon, Methanococcus  
 jannaschii.";  
 RL Science 273:1058-1073(1996).  
 -!- FUNCTION: Catalyzes the reduction of F420 with NADP(+) and the

CC reduction of NADP(+) with F420H(2). Probably couples the NADP-  
 CC dependent oxidation of the alcohol to the aldehyde with the F420-  
 CC dependent reduction of CO(2) to methane (anaerobic function) (By  
 CC similarity).  
 CC -!- CATALYTIC ACTIVITY: NADPH + coenzyme F420 = NADP(+) + reduced  
 CC coenzyme F420.  
 CC -!- SIMILARITY: TO M.ORGANOHILOM F420-DEPENDENT NADP REDUCTASE.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL: U67591; AAB9514.1; -  
 DR PIR: D64487; D64487.  
 DR TIGR: MJ1501; -  
 DR InterPro: IPR004455; NADBoxred\_F420.  
 DR Pfam: PF03807; F420\_oxidored\_1.  
 DR TIGRPFam: TIGR00301; TIGR00301; 1.  
 KW Hypothetical protein; Oxidoreductase; NADP; Complete proteome.  
 SQ SEQUENCE 223 AA; 24068 MW; 2370BDB6F5BBD3D9 CRC64;

Query Match 5.7%; Score 133; DB 1; Length 223;  
 Best Local Similarity 21.9%; Pred. No. 0 003;  
 Matches 46; Conservative 52; Mismatches 78; Indels 34; Gaps 7;

QY 37 GSGDFPKSLTIRLNGYHVIGSNPKPASEFFPHVDV-----THHEDALT 84  
 DB 8 GTDQDFGLALRLAK-NKKIIIGSRKKEAKAKKAEILKRGLEADIGLENDAAK 66  
 QY 85 KTHIIVAIHREH-YTSLMDIRHLVYKTLIDV-----SNNRINQYEP-SNAEVL 133  
 DB 67 EGVVTLISLPYETLSTIQLEKELKTVSGVPLATAGCKPTRLFPDGSVAEAV 126  
 QY 134 ASLPPLSLYKGFNVSAVALQGPDAASROYVIGSNITQASQVIELARQNLFI-PIDL 192  
 DB 127 QNVLSKSKVSAQNVCHAVLELDNPVDCDILVCGNDEBAKKVVIDLANQIDGVADIC 186  
 QY 193 GSLSARAEIENPLRLFTLRGPRVVAISL 222  
 DB 187 GNLEKSRITIEAIT-----PLIGLNI 207

RESULT 3  
 FARE METHH STANDARD; PRT; 232 AA.  
 ID FARE METHH STANDARD; PRT; 232 AA.  
 AC 026350;  
 DT 15-DEC-1998 (Rel. 37, Last Created)  
 DT 15-DEC-1998 (Rel. 37, Last sequence update)  
 DT 10-OCT-2003 (Rel. 42, Last annotation update)  
 DE Putative F420-dependent NADP reductase (EC 1.-.-.-).  
 GN MTH248.  
 OS Methanobacterium thermoautotrophicum.  
 OC Archaea; Euryarchaeota; Methanobacteria; Methanobacteriales;  
 OC Methanobacteriaceae; Methanobacter.  
 OX NCBI\_TaxId=187420;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Delta H;  
 RX MEDLINE=98037514; PubMed=9371463;  
 RA Smith D.R., Doncelte-Stamm L.A., Deloughery C., Lee H.-M., Dubois J.,  
 RA Aldredge T., Bashirzadeh R., Blakey D., Cook R., Gilbert K.,  
 RA Harrison R., Hoang L., Keagle P., Lum W., Pothier B., Qiu D.,  
 RA Spadafora R., Vicare R., Wang Y., Wierzbowski J., Gibson R.,  
 RA Jiwani N., Caruso A., Bush D., Safer H., Patwell D., Prabhakar S.,  
 RA McDougall S., Shiner G., Goyal A., Pietrowski S., Church G.M.,  
 RA Daniels C.J., Mao U.-I., Rice P., Noelling J., Reeve J.N.,  
 RT "Complete genome sequence of Methanobacterium thermoautotrophicum  
 Delta H: functional analysis and comparative genomics.";  
 RL J. Bacteriol. 179:7135-7155(1997).

```

CC -1- FUNCTION: Catalyzes the reduction of F420 with NADP(+) and the
CC reduction of NADP(+) with F420H(2). Probably couples the NADP-
CC dependent oxidation of the alcohol to the aldehyde with the F420-
CC dependent reduction of CO(2) to methane (anabolic function) (By
CC similarity).
CC -1- CATALYTIC ACTIVITY: NADPH + coenzyme F420 = NADP(+) + reduced
CC coenzyme F420.
CC -1- SIMILARITY: TO M.ORGANOPHILUM F420-DEPENDENT NADP REDUCTASE.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch.)
CC -----
DR EMBL; AE000811; AAB84754.1; -.
DR PIR; A69131; A69131.
DR InterPro; IPR004455; MADPOXRED_F420.
DR Pfam; PF03807; F420_oxidored; 1.
DR TIGSFAMS; TIGR00301; TIGR00301; 1.
DR KW Hypothetical protein; Oxidoreductase; NADP; Complete proteome.
SQ SEQUENCE 232 AA; 24539 MW; A1CE60ABC8474236 CRC64;

Query Match 5.1%; Score 120.5; DB 1; Length 232;
Best Local Similarity 23.8%; Pred. No. 0.029;
Matches 49; Conservative 42; Mismatches 90; Indels 25; Gaps 6;

QY 30 KATVGVITSGDPASLTIRLCRGHYVIGSRNKFSEFPFHYDVTHNE----- 80
DB 8 KIAV-IGGTGGQGLAFPAVAGEEYVIGSRDEKSKASKVLELAGRDIDISVEGANT 66
QY 81 -DALTNTNIFVAIH-REHYTSLMDLRLHLVGLKILIVS-----NNMRINQVPSN 129
DB 67 PDAAASADVVLTPPLAQWTLASIRHQVADKTLIDATVPIDSCIGSAVRIYDLMEGS 126
QY 130 AEYLASLF--PDSLIYKGVNVASAMALQLGSPKASQVYICSNNIQARQVIELAQIN 186
DB 127 AAEBAARLEBQGRVVAAPFNINISASALLEVSEFVDCIIVASPHDALVAAELAKID 186
QY 187 FI-PIDGLSSASAREIENLPLRLFTL 211
DB 187 GVRAIECGGLENAKILIKITPLINL 212

RESULT 4
FSHR_MACFA STANDARD; PRT; 695 AA.
ID FSHR_MACFA P32212;
AC 01-OCT-1993 (Rel. 27, Created)
DT 01-OCT-1993 (Rel. 27, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Follicle stimulating hormone receptor precursor (FSH-R) (Follictrophin
DE receptor).
OS Macaca fascicularis (Crab eating macaque) (Cynomolgus monkey).
CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
CC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
CC Cercopithecinae; Macaca.
CC NCBI_TaxID=9541;
OX [1]
RN SEQUENCE FROM N.A.
RP TISSUE=Testis;
RX MEDLINE=94071854; PubMed=7504463.
RA Gromoll U., Danbar B., Sharma R.S., Nieschlag E.;
RT "Molecular cloning of the testicular follicle stimulating hormone
RT receptor of the non human primate Macaca fascicularis and
RT identification of multiple transcripts in the testis.";
RL Biochem. Biophys. Res. Commun. 196;1066-1072(1993).
CC -1- FUNCTION: Receptor for follicle stimulating hormone. The activity
CC of this receptor is mediated by G proteins which activate
CC adenylate cyclase.

```

```

CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
CC -1- SIMILARITY: Belongs to family 1 of G-protein coupled receptors.
CC FSH/LSH/TSH subfamily.
CC -1- SIMILARITY: Contains 6 leucine-rich (LRR) repeats.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch.)
CC -----
DR EMBL; X74454; CA452463.1; -.
DR PIR; JN0898; JN0898.
DR HSSP; P23945; IXUN.
DR InterPro; IPR000276; GPCR_Rhodopsn.
DR InterPro; IPR001611; LRR.
DR InterPro; IPR000372; LRR_Nterm.
DR Pfam; PF00001; 7tm_1; 1.
DR Pfam; PF00560; LRR_3.
DR Pfam; P201462; LRRNT; 1.
DR PRINTS; PR00237; GPCR_RHODOPSN.
DR SMART; SM00013; LRRNT; 1.
DR PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
DR PROSITE; PS0262; G_PROTEIN_RECEP_F1_2; 1.
DR G-protein coupled receptor; Transmembrane; Glycoprotein; Signal;
KW Phosphorylation; Repeat; Leucine-rich repeat.
FT SIGNAL 1
FT CHAIN 18
FT DOMAIN 18 695
FT TRANSMEM 367 387
FT DOMAIN 388 398
FT TRANSMEM 399 421
FT DOMAIN 422 443
FT TRANSMEM 444 465
FT DOMAIN 466 485
FT TRANSMEM 486 508
FT DOMAIN 509 528
FT TRANSMEM 529 550
FT DOMAIN 551 573
FT TRANSMEM 574 597
FT DOMAIN 598 608
FT TRANSMEM 609 630
FT DOMAIN 631 695
FT REPEAT 44 68
FT REPEAT 69 93
FT REPEAT 119 143
FT REPEAT 170 192
FT REPEAT 193 216
FT REPEAT 218 240
FT DISULFID 442 517
FT CARBOHYD 191 191
FT CARBOHYD 199 199
FT CARBOHYD 293 293
FT CARBOHYD 318 318
SQ SEQUENCE 695 AA; 78343 MW; 0D60A233729B5250 CRC64;

Query Match 4.7%; Score 111.5; DB 1; Length 695;
Best Local Similarity 17.0%; Pred. No. 0.56;
Matches 84; Conservative 75; Mismatches 151; Indels 183; Gaps 16;

QY 79 REDALTNTNIFVAIHREHYTSLMDLRLVLG-----KILIDVSNMR 121
DB 98 HEIRIEKANML-LYNPEAFQNLPLRLTLISNTGIKILPDVHKHSFOKVLIDIDQVIN 156
QY 122 INQVESNAEYIASLFPDSLIVKGVNVASAMALQLGPDASROYICSNNIQARQVIEL 181
DB 157 IH-----TIERSFVGLSFESYIIML-----KNGGQEHANCA----- 189
QY 182 AQLNFIPIIDGLSSASAREIENLPLRLFTLRGSPVVAISLATFPFLYSFVRDIHPYAR 241
DB 190 ---FNGTQIDELNLSDDNNLDELPLNDVFHGASGVIIIDISRTIRHSFPSYGLNMLKCLR 246

```

```

QY 242 NQOSDFYKPIEIVNKLPIVAILSLVY----- 271
DB 247 RSTVNLKLP-----SLERKVALMEASITYPHSCAFANMRROISELHPICNKSILRQEV 301
QY 272 -----LAGLAAAYGLVYGTXYRRF-----PPMLEWTWLCRRQLG 306
DB 302 DYMOTRQGRSSLAEDNESSISRGFDMTYAFDLDCEVVDVTCSPDPANFDCEDILG 361
QY 307 -----LISFPAMVHA-----YSLCLPMRSEERYFLNMAVOOVANIENS 348
DB 362 YNLEVLIMFISILAITGNITVLTLSQYKLTVP-----RFLMCULAFAD----- 408
QY 349 WNEEVRWRIEMVYISGIMSLGSLIANTSIPIVSNALNME-----FSFIQSTLG 399
DB 409 -----LCIGYVLLIASVDIHTKSOYHNYALDWQTAGAGCDAGCFVYFASELS 456
QY 400 ---YVALLISTFH-----VLIGMKRAFEERYRFTPPNFV-LAL 436
DB 457 VYTLAITLERMHTTHAMQDCKVHVAASVWVMGRIFAALAPPIRQISSYMKVSI 516
QY 437 VLPSTVILDLQL 449
DB 517 CLPMDIDPSLSQL 529

```

## RESULT 5

Y538\_PASMTU STANDARD; PRT; 206 AA.

```

AC Q9C997;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Hypothetical protein PM0538.
GN PM0538.
OS Pasteurella multocida.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
OC Pasteurellaceae; Pasteurella.
OX NCBI_TaxID=747;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=PM70;
RX MEDLINE=21145866; PubMed=11248100;
RA May B.-J., Zhang Q., Li L.-L., Paustian M.L., Whitem T.S., Kapur V.;
RT "Complete genomic sequence of Pasteurella multocida PM70.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:3460-3465(2001).
CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
CC -1- SIMILARITY: Belongs to the UPF0191 family.

```

-----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----

```

DR EMBL; AE006089; AAK02622.1; -
DR HAMAP; MF_01207; -; 1.
DR InterPro; IPR007916; UPF0191.
DR Pfam; PF05252; UPF0191; 1.
KW Hypothetical protein; Transmembrane; Complete proteome.
FT TRANSMM 7
FT TRANSMM 26
FT TRANSMM 41
FT TRANSMM 63
FT TRANSMM 76
FT TRANSMM 98
FT TRANSMM 113
FT TRANSMM 135
FT TRANSMM 142
FT TRANSMM 164
FT TRANSMM 169
FT TRANSMM 191
SQ SEQUENCE 206 AA; 22F77729D7F11194 CRC64;

```

Query Match 4.7%; Score 109.5; DB 1; Length 206;  
 Best Local Similarity 23.8%; Pred. No. 0.18;  
 Matches 56; Conservative 44; Mismatches 72; Indels 63; Gaps 13;

```

QY 228 LYSFVADVTH-----PYA-----NQOSDFYKPIEIVNKLPIVAILSLVYLAGL 275
DB 1 MSLFRIILIHVOCGLGVAMVLWSGDSGADPIKEIHQFLGFSALITLIMFIIG- 59
QY 276 LAAAYGLVYGTXYRRFPPLLEWTWLCRRQLGLSSFPAMVHAAYSLCLPMRSEERYFLN 335
DB 60 -----KVFFLYLK-----QPOLQV-----LRRLGLMAMFYVVAHV-YA-----YLALE 96
QY 336 MAYQOVHANIENSWNEEVRWRIEMVYISGIMSLGSLIANTSIPIVSNALNMEFSFIQ 395
DB 97 LGVD-----FSLFQDELNR-----GYLLIGALIFLITLMAALSSWGYLK-----LKGKMWFTL 146
DB 147 HOLGYVALLIGAIH---YVM-----SVKNVTFSSML---YLILSIMILC 184

```

## RESULT 6

FSHR\_CHICK STANDARD; PRT; 693 AA.

```

AC F79763; Q90719;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Follicle stimulating hormone receptor precursor (FSH-R) (Folliotropin  

  DE receptor).
GN FSHR.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  

OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;  

OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Ovary;
RX MEDLINE=97473503; PubMed=9332357;
RA Wakabayashi N., Suzuki A., Hoshino H., Nishimori K., Mizuno S.;  

RT "The cDNA cloning and transient expression of a chicken gene encoding  

  RT a follicle-stimulating hormone receptor.";  

RL Gene 197:121-127(1997).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=White Leghorn;
RX MEDLINE=97057887; PubMed=8902217;
RA You S., Bridgman J.T., Foster D.N., Johnson A.L.;  

RT "Characterization of the chicken follicle stimulating hormone  

  RT receptor (FSH-R) complementary deoxyribonucleic acid, and expression  

  RT of cFSH-R messenger ribonucleic acid in the ovary.";  

RL Biol. Reprod. 55:1055-1062(1996).
CC -1- FUNCTION: Receptor for follicle stimulating hormone. The activity  

  CC of this receptor is mediated by G proteins which activate  

  CC adenylyl cyclase.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
CC -1- SIMILARITY: Belongs to family 1 of G-protein coupled receptors.
CC FSH/LSH/FSH subfamily.
CC -1- SIMILARITY: Contains 7 leucine-rich (LRR) repeats.

```

-----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----

```

DR EMBL; D87871; BAAL3487.1; -
DR HBBP; U51097; AACG0030.1; -
DR HSSP; P23945; LXUN.
DR InterPro; IPR000276; GPCR_Rhodopsn.
DR InterPro; IPR001611; LRR.
DR InterPro; IPR000372; LRR_Nterm.
DR Pfam; PF00001; 7cm_1; 1.

```

DR Pfam: PF00560; LRR; 2.  
 DR PRINTS: SM00237; GPCRHRDOPSN.  
 DR SMART: SM00013; LRNT; 1.  
 DR PROSITE: PS00237; G-PROTEIN-RECEP\_F1\_1; 1.  
 DR PROSITE: PS50262; G-PROTEIN-RECEP\_F1\_2; 1.  
 KW G-protein coupled receptor; Transmembrane; Glycoprotein; Signal;  
 KW Phosphorylation; Repeat; Leucine-rich repeat.  
 FT SIGNAL 1 17  
 FT CHAIN 18 693  
 FT DOMAIN 18 366  
 FT TRANSMEM 367 387  
 FT DOMAIN 388 398  
 FT TRANSMEM 399 421  
 FT DOMAIN 422 443  
 FT TRANSMEM 444 465  
 FT DOMAIN 466 485  
 FT TRANSMEM 486 508  
 FT DOMAIN 509 528  
 FT TRANSMEM 529 550  
 FT TRANSMEM 551 573  
 FT TRANSMEM 574 597  
 FT TRANSMEM 598 608  
 FT TRANSMEM 609 630  
 FT DOMAIN 631 693  
 FT REPEAT 45 68  
 FT REPEAT 69 93  
 FT REPEAT 95 118  
 FT REPEAT 119 143  
 FT REPEAT 168 192  
 FT REPEAT 193 216  
 FT REPEAT 218 240  
 FT DISULFID 442 517  
 FT CARBOHYD 47 47  
 FT CARBOHYD 191 191  
 FT CARBOHYD 199 199  
 FT CARBOHYD 268 268  
 FT CONFLICT 4 4  
 FT CONFLICT 88 88  
 FT CONFLICT 140 140  
 FT CONFLICT 174 174  
 FT CONFLICT 191 191  
 FT CONFLICT 329 329  
 SQ SEQUENCE 693 AA; 78697 MW; 46F9869635A1BEC CRC64;  
 Query Match 4.6%; Score 109; DB 1; Length 693;  
 Best Local Similarity 17.2%; Pred. No. 0.87;  
 Matches 100; Conservative 72; Mismatches 173; Indels 238; Gaps 21;  
 QY 18 LPNG-INGIKDARKVTVG---VIGSGDFAKSITRLIRGCHVAVIGSNPKPASEFP 71  
 DB 61 IPKAFGLHLEKIEISQNDALBIEIGNVFSS- PKL----- 97  
 QY 72 HAVDVTHEDALTNTIIFVAIHREHYTSAMDRLHVLVG-----KILI 114  
 DB 98 -----HEIRIEKANKL-MKIDDAFOHPLSLVYLISNTGSLFVYKAKHSFOKVL 149  
 QY 115 DVSNMNRINQYPSNAEYLAFLPDSLIVKGFNVASAMALQGPKASRGVYICSNNTQA 174  
 DB 150 DVQNIHRIETERTFWGLSS--ESVYLR-----LNKQIOE 184  
 QY 175 RQVIELARQLNFPIDUGLSASARETENPLRFTLMRGVVAISLAFEPFLYSVRD 234  
 DB 185 IKD-----HAFNGTCLBELNLSDNVNEKPEKFGQAIGVVDISRTISFLPSHGLE 239  
 QY 235 VVHVPANQOSDFPKIPLEIVNKTLPVATLLSLVIAGL----- 275  
 DB 240 FIKKLARSYKLLKLP--DVNKRSLIEANFTVPSHCAFTNRKQNTSEFYPICSMPA 297  
 QY 276 -----LAAAYQLYGTRY-----RRF-----PPMLETWLOCR 302  
 DB 298 KQDIEQTGKKHRRSAADYISHYGTRFGPVENEFPYGLCNEVVDVCSPKDAPNCE 357  
 QY 303 KQLG-----LISFFAMVHVA-----YSCLPMRSEERYLFANVAYQVHAN 344

DB 358 DINGYNVLRVLIWFINILATIGNTVTLILISSQYLTVP-----RFLMGNLAFAD----- 408  
 QY 345 IENSNNEEHWREIEMYSFGISLGLSLAVTSIPSVSANALNRE-----PSFIQ 395  
 DB 409 -----LCIGYILFLASVDIQTKSYNVYALDMQAGACNNAAGFTTVA 452  
 QY 396 STIGYALLISFPH-----VLIIYKRAF----- 419  
 DB 453 SELSVYTLTVTLIERHTTYAMQINRKVRLNHAIVIMFGWVFATVLLPIFGISSYM 512  
 QY 420 -----EEERYFPTPNPVVALVPSVIIDLOLCRY 452  
 DB 513 KVISCLPMHIEPFSQAYV--IFLVANVLAFTVILICITICY 553  
 RESULT 7  
 ID PRE6 YEAST STANDARD; PRT; 712 AA.  
 AC 012473;  
 DT 15-DEC-1998 (Rel. 37, Created)  
 DT 15-DEC-1998 (Rel. 37, Last sequence update)  
 DT 10-OCT-2003 (Rel. 42, Last annotation update)  
 DE Ferric reductase transmembrane component 6 precursor (EC 1.16.1.7)  
 DE (ferric-chelate reductase 6).  
 GN PRE6 OR YIL051C OR I0593.  
 OS Saccharomyces cerevisiae (Baker's yeast).  
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
 OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.  
 NC NCBI Taxid=4932;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=8288C / AB972;  
 RX MEDLINE=97313267; PubMed=9169871;  
 RA Johnston M., Hillier L., Riles L., Albermann K., Andre B., Ansoorge W.,  
 RA Benes V., Bruecker M., Delius H., Dubois E., Duesterhoeft A.,  
 RA Entian K.-D., Floeth M., Goffeau A., Hebling U., Heumann K.,  
 RA Heuss-Neitzel D., Hilbert H., Hilger F., Kleine K., Koester P.,  
 RA Louis E.J., Messerguy F., Mewes H.-W., Miesga T., Moestl D.,  
 RA Mueller-Auer S., Neuwirth U., Obermaier B., Piravandi B., Pohl T.M.,  
 RA Portetelle D., Purnelle B., Reichmann S., Riesger M., Rinke M., Rose M.,  
 RA Scharte M., Scherens B., Scholler P., Schwager C., Schwarz S.,  
 RA Underwood A.P., Urrutiarazu L.A., Vandenbol M., Verhaesselt P.,  
 RA Vierendeels F., Voet M., Volckaert G., Voss H., Wambutt R., Medler E.,  
 RA Medler H., Zimmermann F.K., Zollner A., Hant J., Hoheisel J.D.;  
 RT "The nucleotide sequence of Saccharomyces cerevisiae chromosome XII.";  
 RL Nature 387:87-90(1997).  
 CC -1- CATALYTIC ACTIVITY: NADH + 2 Fe(3+) = NAD(+) + 2 Fe(2+).  
 CC -1- COFACTOR: FAD (Probable).  
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein (potential).  
 CC -1- SIMILARITY: Belongs to the PRE / CYB8 family.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@sib.ch](mailto:license@sib.ch)).  
 CC  
 DR EMBL: Z47973; CAAB8006.1; -  
 DR EMBL: Z73156; CAA97503.1; -  
 DR PIR: S50969; S50969.  
 DR GeneID: 142046; -  
 DR SGD: S0003974; PRE6.  
 DR InterPro: IPR002916; Ferric\_reduct.  
 DR Pfam: PF01794; Ferric\_reduct\_1.  
 KW Oxidoreductase; Electron transport; Transmembrane; Iron transport;  
 KW FAD; NAD; Glycoprotein; Signal; Multigene family.  
 FT SIGNAL 1 17  
 FT CHAIN 18 712  
 FT NP\_BIND 493 499  
 FT FAD (POTENTIAL). 6.

```

FT TRANSMEM 168 168 POTENTIAL.
FT TRANSMEM 245 265 POTENTIAL.
FT TRANSMEM 308 308 POTENTIAL.
FT TRANSMEM 329 349 POTENTIAL.
FT TRANSMEM 361 381 POTENTIAL.
FT TRANSMEM 388 408 POTENTIAL.
FT TRANSMEM 417 437 POTENTIAL.
FT TRANSMEM 478 498 POTENTIAL.
FT TRANSMEM 550 570 POTENTIAL.
FT DOMAIN 553 556 POLY-LEU.
FT CARBOHYD 89 89 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 112 112 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 124 124 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 450 450 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 655 655 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 712 AA; 81989 MW; 5224F12B51544BAA CRC64;

Query Match 4.6%; Score 109; DB 1; Length 712;
Best Local Similarity 18.8%; Pred. No. 0.9; Indels 142; Gaps 17;
Matches 72; Conservative 51; Mismatches 119;

QY 88 IIFVALHREHYTSL-----WDLHLLVGKILIDVSNMKNINQYPSNAEYASL 136
DB 183 IIAVFPMHSHYNGLANRALFASRFVYIRGHFVLPFLVD-----KHANHEKFLNVEFTGL 238
QY 137 FPDLSLVKGFNVSAHALQGPDAKROYICSNNOIQAQOYIELARQINFIDLSLS 196
DB 239 MPNSL-----EAM----- 246
QY 197 SAREIENLPLRLFTLRGPVVAVALSATEFFL-YSFQVDVHPYARNOQSPFYKPIBIV 255
DB 247 -----IIRGYTLANTIFLISISYIIDPYNLFINSHLSQTRL--LA 284
QY 256 NKTLPVALTTLSTVLAGILAAAYOLYGTKRPPPLETWLOCRKQGLSFPFAMV 315
DB 285 DRS-GILAFQFPLIIIFTRNSFLFLTGKNSF-----ISFKMIGIMVLANAT 336
QY 316 H-AVASYICLPMRSEBYLFLNMAVQOVHANIEENSWEEWRIEMTISGIMSLGLSL 374
DB 337 HSLSYSL-----FAINNAFK-----ISNK-----QLYWKSGISITVLCVL 373
QY 375 AVTSIPSVSNALNREBSFIOSTLGYVALII--STFHVLIY-GWKAFE-----EE 422
DB 374 LVLSLGLVGRK---RAYEEFLYTHIITALLFFYCCQWYKIFNGWEMVLSLLIWGLEK 429
QY 423 YRFPYPPNFVALVPSIVILDL 446
DB 430 LFRIMW---ILOFRFPKATLINL 449

RESULT 8
PRE7 YEAST ID STANDARD; PRT; 629 AA.
AC Q1233;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Ferric reductase transmembrane component 7 (BC 1.16.1.7) (Ferric-
DE chelate reductase 7).
GN PRE7 OR YOL152W.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
RX SEQUENCE FROM N.A.
RC STRAIN=8288C / FY1679;
RX MEDLINE=96132030; PubMed=8553699;
RA Casamayor A., Aldea M., Casas C., Herrero E., Gano F.J.,
RA Launente M.J., Gancedo C., Arino J.,
RT "DNA sequence analysis of a 13 kbp fragment of the left arm of yeast
RT chromosome XV containing seven new open reading frames.";
RL Yeast 11:1281-1288(1995).

```

```

CC -1- CATALYTIC ACTIVITY: NADH + 2 Fe(3+) = NAD(+) + 2 Fe(2+).
CC -1- COFACTOR: FAD (Probable).
CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
CC -1- SIMILARITY: Belongs to the FRE / CYB5 family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; Z48239; CA88276.1; -.
DR EMBL; Z74894; CA99174.1; -.
DR PIR; S60385; S60385.
DR SCDB; S0005512; PRE7.
DR InterPro; IPR002916; Ferric reduct.
DR Pfam; PF01794; Ferric reduct. 1.
KM Oxidoreductase, Electron transport; Transmembrane; Iron transport;
KM FAD; NAD; Glycoprotein; Multigene family.
KM NP_BIND; 369 375 FAD (POTENTIAL).
FT TRANSMEM 46 66 POTENTIAL.
FT TRANSMEM 108 128 POTENTIAL.
FT TRANSMEM 168 188 POTENTIAL.
FT TRANSMEM 195 215 POTENTIAL.
FT TRANSMEM 238 258 POTENTIAL.
FT TRANSMEM 266 286 POTENTIAL.
FT TRANSMEM 293 313 POTENTIAL.
FT TRANSMEM 422 442 POTENTIAL.
FT CARBOHYD 330 330 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 541 541 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 629 AA; 71996 MW; 2384480E9289C16F CRC64;

Query Match 4.6%; Score 107.5; DB 1; Length 629;
Best Local Similarity 21.7%; Pred. No. 1;
Matches 57; Conservative 46; Mismatches 77; Indels 83; Gaps 17;

QY 214 GPVVVALSLATFFLYFVVDVHPYARNOQSPFYKPIBIVNKTLPVALTTLSTVY-L 272
DB 123 GTFVYVMAITLTYLLYCFV---HPFYR-PCAEGSPPLSV---RAGIMASLVPFVSL 175
QY 273 AG-----LIAAYQOLYGTKRPPPLETWLOCRKQGLSFPFAMVAYSLCLPM 325
DB 176 SGKINVGMLVGLSYE---KINITYHW-----ASIICTFSPSWHV-----IPF 215
QY 326 RRSERYLFLNMAVQOVHANIEENSWEEWRIEMTISFGIMSLGLSLAVTSIP----- 380
DB 216 LRQAKH---EGYERKH---QWKASDWRK-----SSVPIFLFLNLWLSLPIARRH 262
QY 381 --SVSNALNREBSFIOSTLGYVALIISTFHV-----LIYWKRAFEERYR- 425
DB 263 FYEIFLQHW-----LIAVGYSILF---YHVPDLNSHWLVATIVVW---PAQLFYRL 311
QY 426 ----FYTPPNFVALVPSIVIL 444
DB 312 AVKGYLRPGSGFASSTANVSTV 334

RESULT 9
FSHR HUMAN ID STANDARD; PRT; 695 AA.
AC P23945; Q16225;
DT 01-MAR-1992 (Rel. 21, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Folicle stimulating hormone receptor precursor (FSH-R) (Folllitropin
DE receptor).
GN FSHR.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

```



OK NCBI\_Taxid=9606;  
 RN [1]  
 RN SEQUENCE FROM N.A., AND VARIANTS ALA-307 AND SER-680.  
 RC TISSUE=Ovary;  
 RA MEDLINE=91222171; PubMed=1709010;  
 RA Miregish T., Nakamura K., Takakura Y., Ihuki Y., Igarashi M.,  
 RT "Cloning and sequencing of human FSH receptor cDNA,"  
 RL Biochem. Biophys. Res. Commun. 175:1125-1130(1991).  
 RN [2]  
 RN SEQUENCE FROM N.A.  
 RC TISSUE=Testis;  
 RA MEDLINE=93246012; PubMed=1301382;  
 RA Kelton C.A., Cheng S.V., Nugent N.P., Schweickhardt R.L.,  
 RA Rosenthal J.L., Overton S.A., Wands G.D., Kuzela J.B., Luchette C.A.,  
 RA Chapel S.C.,  
 RT "The cloning of the human follicle stimulating hormone receptor and  
 its expression in COS-7, CHO, and Y-1 cells,"  
 RL Mol. Cell. Endocrinol. 89:141-151(1992).  
 RN [3]  
 RN SEQUENCE FROM N.A.  
 RA Tilly L.T., Alhara T., Nishimori K., Jai X.-C., Billig H.,  
 RA Kowalski K.I., Perlas E.A., Hsueh A.O.,  
 RL Submitted (XX-1992) to the EMBL/GenBank/DBJ databases.  
 RN [4]  
 RN SEQUENCE OF 1-342 FROM N.A.  
 RC TISSUE=Testis;  
 RA MEDLINE=93075197; PubMed=1359889;  
 RA Gromoll U., Gudermann T., Nieschlag E.,  
 RT "Molecular cloning of a truncated isoform of the human follicle  
 stimulating hormone receptor,"  
 RL Biochem. Biophys. Res. Commun. 188:1077-1083(1992).  
 RN [5]  
 RN SEQUENCE OF 1-51 FROM N.A.  
 RA MEDLINE=95011044; PubMed=7926278;  
 RA Gromoll U., Dankhar B., Gudermann T.,  
 RT "Characterization of the 5' flanking region of the human follicle-  
 stimulating hormone receptor gene,"  
 RL Mol. Cell. Endocrinol. 102:93-102(1994).  
 RN [6]  
 RN SEQUENCE OF 286-695 FROM N.A., AND VARIANT SER-680.  
 RA MEDLINE=95000244; PubMed=7916967;  
 RA Gromoll U., Red T., Holtgreve-Grez H., Nieschlag E., Gudermann T.,  
 RT "Localization of the human FSH receptor to chromosome 2p21 using a  
 genomic probe comprising exon 10,"  
 RL J. Mol. Endocrinol. 12:265-271(1994).  
 RN [7]  
 RN 3D-STRUCTURE MODELING OF 48-228.  
 RA MEDLINE=96363672; PubMed=8747461;  
 RA Jiang X., Dreano M., Buckler D.R., Cheng S., Ythier A., Wu H.,  
 RA Hendrickson W.A., el Tayar N.,  
 RT "Structural predictions for the ligand-binding region of glycoprotein  
 hormone receptors and the nature of hormone-receptor interactions,"  
 RL Structure 3:141-153(1995).  
 RN [8]  
 RN VARIANTS ALA-307, ARG-524 AND SER-680.  
 RA MEDLINE=99318093; PubMed=10391209;  
 RA Cargill M., Altschuler D., Ireland J., Sklar P., Ardlie K., Patil N.,  
 RA Shaw N., Lane C.R., Lim E.P., Kalyanaraman N., Nimesh J., Ziaugra L.,  
 RA Friedland L., Rolfe A., Warrington J., Lipshutz R., Daley G.Q.,  
 RA Lander E.S.,  
 RT "Characterization of single-nucleotide polymorphisms in coding regions  
 of human genes,"  
 RL Nat. Genet. 22:231-238(1999).  
 RN [9]  
 RN ERRATUM.  
 RA Cargill M., Altschuler D., Ireland J., Sklar P., Ardlie K., Patil N.,  
 RA Shaw N., Lane C.R., Lim E.P., Kalyanaraman N., Nimesh J., Ziaugra L.,  
 RA Friedland L., Rolfe A., Warrington J., Lipshutz R., Daley G.Q.,  
 RA Lander E.S.,  
 RL Nat. Genet. 23:373-373(1999).  
 RN [10]  
 RN VARIANTS ALA-307 AND SER-680.  
 RA MEDLINE=22054685; PubMed=12059813;

RA Aastiani K., Gromoll U., Eckardtstein S.V., Zitzmann M., Nieschlag E.,  
 RA Simoni M.,  
 RT "Distribution and function of FSH receptor genetic variants in normal  
 men,"  
 RL Andrologia 34:172-176(2002).  
 RN [11]  
 RN VARIANT OHSS IIE-449.  
 RA MEDLINE=22812036; PubMed=12930927;  
 RA Vasseur C., Rodien P., Beau I., Desroches A., Gerard C.,  
 RA de Poncheville L., Chaplot S., Savagner F., Croue A., Mathieu E.,  
 RA Lahou N., Descamps P., Mirahil M.,  
 RT "A chorionic gonadotropin-sensitive mutation in the  
 follicle-stimulating hormone receptor as a cause of familial  
 gestational spontaneous ovarian hyperstimulation syndrome,"  
 RL New Engl. J. Med. 349:753-759(2003).  
 RN [12]  
 RN VARIANT OHSS ASN-567.  
 RA MEDLINE=22812037; PubMed=12930928;  
 RA Smits G., Olatunbosun O., Delbaere A., Pierson R., Vassart G.,  
 RA Costagliola S.,  
 RT "Ovarian hyperstimulation syndrome due to a mutation in the  
 follicle-stimulating hormone receptor,"  
 RL New Engl. J. Med. 349:760-766(2003).  
 CC -1- FUNCTION: Receptor for follicle stimulating hormone. The activity  
 of this receptor is mediated by G proteins which activate  
 adenylyl cyclase.  
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein.  
 CC -1- ALTERNATIVE PRODUCTS:  
 CC Event=Alternative splicing; Named Isoforms=2;  
 CC Name=Long;  
 CC IsoId=P23945-1; Sequence=Displayed;  
 CC Name=Short;  
 CC IsoId=P23945-2; Sequence=VSP 001953, VSP 001954;  
 CC -1- TISSUE SPECIFICITY: Sertoli cells and ovarian granulosa cells.  
 CC -1- DISEASE: Defects in FSHR are a cause of ovarian hyperstimulation  
 syndrome (OHSS) [MIM:608115]. OHSS is a disorder which occurs  
 either spontaneously or most often as an iatrogenic complication  
 of ovarian stimulation treatments for in vitro fertilization. The  
 clinical manifestations vary from abdominal distention and  
 discomfort to potentially life-threatening, massive ovarian  
 enlargement and capillary leak with fluid sequestration.  
 CC Pathologic features of this syndrome include the presence of  
 multiple serous and hemorrhagic follicular cysts lined by  
 luteinized cells, a condition called hyperreactic luteinels.  
 CC -1- SIMILARITY: Belongs to family 1 of G-protein coupled receptors.  
 CC FSH/LSH/FSH subfamily.  
 CC -1- SIMILARITY: Contains 5 leucine-rich (LRR) repeats.  
 CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC EMBL: M65085; AAA52477.1; -  
 DR EMBL: SS9900; AAB26480.1; -  
 DR EMBL: M95489; AAB2478.1; -  
 DR EMBL: X68044; CAA48179.1; -  
 DR EMBL: S73199; AAB32071.1; -  
 DR EMBL: S73526; AAB32225.1; -  
 DR PIR: I57661; ORHDFT.  
 DR PDB: 1XUN; 15-MAY-97.  
 DR Genew; HGNC:3969; FSHR.  
 DR MIM: 136435; -  
 DR MIM: 608115; -  
 DR GO: GO:0016021; C:Integral to membrane; TAS.  
 DR GO: GO:0005886; C:Plasma membrane; TAS.  
 DR GO: GO:0004963; F:follicle stimulating hormone receptor activity; TAS.  
 DR GO: GO:0007292; P:female gamete generation; TAS.  
 DR GO: GO:0006885; P:female gonad development; TAS.  
 DR GO: GO:0007186; P:G-protein coupled receptor protein signalin. .; TAS.

DR GO; GO:0007283; P:spermatogenesis; TAS.  
 DR InterPro; IPR000276; GPCR\_Rhodopsn.  
 DR InterPro; IPR001611; LRR.  
 DR InterPro; IPR000372; LRR\_Nterm.  
 DR Pfam; PF00001; 7tm\_1; 1.  
 DR Pfam; PF00560; LRR; 3.  
 DR Pfam; PF01462; LRNT; 1.  
 DR PRINTS; PR00237; GPCR\_Rhodopsn.  
 DR SMART; SM00013; LRNT; 1.  
 DR PROSITE; PS00237; G\_PROTEIN\_RECEP\_F1\_1; 1.  
 DR PROSITE; PS0262; G\_PROTEIN\_RECEP\_F1\_2; 1.  
 DR G-protein coupled receptor; Transmembrane; Glycoprotein; Signal;  
 KM Phosphorylation; Repeat; Leucine-rich repeat; Alternative splicing;  
 KW 3D-structure; Polymorphism; Disease mutation.  
 FT SIGNAL 1 17  
 FT CHAIN 1 695  
 FT DOMAIN 18 366  
 FT TRANSMEM 367 387  
 FT DOMAIN 388 398  
 FT TRANSMEM 399 421  
 FT DOMAIN 422 443  
 FT TRANSMEM 444 465  
 FT DOMAIN 466 485  
 FT TRANSMEM 486 508  
 FT DOMAIN 509 528  
 FT TRANSMEM 529 550  
 FT DOMAIN 551 573  
 FT TRANSMEM 574 597  
 FT DOMAIN 598 608  
 FT TRANSMEM 609 630  
 FT DOMAIN 631 695  
 FT REPEAT 69 93  
 FT REPEAT 119 143  
 FT REPEAT 170 192  
 FT REPEAT 193 216  
 FT REPEAT 218 240  
 FT DISULFID 442 517  
 FT CARBOHYD 191 191  
 FT CARBOHYD 199 199  
 FT CARBOHYD 293 293  
 FT CARBOHYD 318 318  
 FT VARSPPLIC 224 285  
 FT /FTId=VSP\_001953.  
 Query Match 4.6%; Score 107.5; DB 1; Length 695;  
 Best Local Similarity 17.0%; Pred. No. 1.1;  
 Matches 84; Conservative 77; Mismatches 148; Indels 185; Gaps 17;

DR 409 -----ICGIYLLIASVDINTKSYQNHAYADWOTGAGCDAAGFTVPASL 455  
 QY 399 G---YVALNIFPR-----VLIYGRKRFEEERYFYFPNPFV-LA 435  
 DB 456 SVYTLNRIETLRIHTHTTHAMQDCKYQLRHAASVYMGVIFAPALAFIFIGISIMKVS 515  
 QY 436 LVIPSTVILLDIQL 449  
 DB 516 ICLPMDIDSPLSQL 529  
 RESULT 10  
 ID FSHR\_PIG STANDARD; PRT; 695 AA.  
 AC P49059; 077514;  
 DT 01-FEB-1996 (Rel. 33, Created)  
 DT 15-DEC-1998 (Rel. 37, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Follicle stimulating hormone receptor precursor (FSH-R) (folliotropin receptor).  
 GN FSHR.  
 OS Sus scrofa (Pig).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.  
 OX NCBI\_TaxID=9823;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Ovary;  
 RX MEDLINE=96011644; PubMed=7590277;  
 RA Remy J.J., Lahbid-Manaas Y., Yerie M., Bozon V., Couture L., Pajot E., Grebert D., Salese R.;  
 RT "The porcine follicle stimulating hormone receptor: cDNA cloning, functional expression and chromosomal localization of the gene";  
 RL Gene 163:257-261 (1995).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Ovary;  
 RA Wang Y.F., Meyer K.B., Schmidt K., Wan S.J., Degen S.J.F., la Barbera A.R.;  
 RT "Porcine follicle-stimulating hormone receptor";  
 RL Submitted (Sep-1997) to the EMBL/GenBank/DBJ databases.  
 CC - FUNCTION: Receptor for follicle stimulating hormone. The activity of this receptor is mediated by G proteins which activate adenylate cyclase.  
 CC - SUBCELLULAR LOCATION: Integral membrane protein.  
 CC - SIMILARITY: Belongs to family 1 of G-protein coupled receptors.  
 CC - SIMILARITY: Contains 6 leucine-rich (LRR) repeats.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 DR EMBL; L1966; AA86933.1; -.  
 DR EMBL; AF025377; AAC24981.1; -.  
 DR HSSP; P23945; 1XUN.  
 DR InterPro; IPR000276; GPCR\_Rhodopsn.  
 DR InterPro; IPR001611; LRR.  
 DR InterPro; IPR000372; LRR\_Nterm.  
 DR Pfam; PF00001; 7tm\_1; 1.  
 DR Pfam; PF00560; LRR; 4.  
 DR Pfam; PF01462; LRNT; 1.  
 DR PRINTS; PR00237; GPCR\_Rhodopsn.  
 DR SMART; SM00013; LRNT; 1.  
 DR PROSITE; PS00237; G\_PROTEIN\_RECEP\_F1\_1; 1.  
 DR PROSITE; PS0262; G\_PROTEIN\_RECEP\_F1\_2; 1.  
 DR G-protein coupled receptor; Transmembrane; Glycoprotein; Signal;  
 KM Phosphorylation; Repeat; Leucine-rich repeat.

FT	SIGNAL	1	17	POTENTIAL.
FT	CHAIN	18	695	POTLICE STIMULATING HORMONE RECEPTOR.
FT	DOMAIN	18	366	EXTRACELLULAR (POTENTIAL).
FT	TRANSSEM	367	387	1 (POTENTIAL).
FT	DOMAIN	388	398	CYTOPASMIC (POTENTIAL).
FT	TRANSSEM	399	421	2 (POTENTIAL).
FT	DOMAIN	422	443	EXTRACELLULAR (POTENTIAL).
FT	TRANSSEM	444	465	3 (POTENTIAL).
FT	DOMAIN	466	485	CYTOPASMIC (POTENTIAL).
FT	TRANSSEM	486	508	4 (POTENTIAL).
FT	DOMAIN	509	528	EXTRACELLULAR (POTENTIAL).
FT	TRANSSEM	529	550	5 (POTENTIAL).
FT	DOMAIN	551	573	CYTOPASMIC (POTENTIAL).
FT	TRANSSEM	574	597	6 (POTENTIAL).
FT	DOMAIN	598	608	EXTRACELLULAR (POTENTIAL).
FT	TRANSSEM	609	630	7 (POTENTIAL).
FT	DOMAIN	631	695	CYTOPASMIC (POTENTIAL).
FT	REPEAT	64	68	LRR 1.
FT	REPEAT	69	93	LRR 2.
FT	REPEAT	119	143	LRR 3.
FT	REPEAT	170	192	LRR 4.
FT	REPEAT	193	216	LRR 5.
FT	REPEAT	218	240	LRR 6.
FT	DISULFID	442	517	BY SIMILARITY.
FT	CARBOHYD	191	191	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	199	199	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	293	293	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CONFLICT	2	2	S -> A (IN REF. 1).
FT	CONFLICT	13	13	T -> S (IN REF. 1).
FT	CONFLICT	60	60	V -> A (IN REF. 1).
FT	CONFLICT	166	166	V -> M (IN REF. 1).
FT	CONFLICT	215	215	Q -> H (IN REF. 1).
FT	CONFLICT	247	247	K -> R (IN REF. 1).
FT	CONFLICT	257	257	S -> T (IN REF. 1).
FT	CONFLICT	334	334	D -> N (IN REF. 1).
FT	CONFLICT	349	349	E -> K (IN REF. 1).
FT	CONFLICT	352	352	T -> A (IN REF. 1).
FT	CONFLICT	383	383	V -> E (IN REF. 1).
FT	CONFLICT	407	407	A -> T (IN REF. 1).
FT	CONFLICT	421	421	V -> I (IN REF. 1).
FT	CONFLICT	427	427	T -> S (IN REF. 1).
FT	CONFLICT	435	435	D -> N (IN REF. 1).
FT	CONFLICT	483	483	L -> V (IN REF. 1).
FT	CONFLICT	550	550	T -> I (IN REF. 1).
FT	CONFLICT	586	586	A -> V (IN REF. 1).
FT	CONFLICT	607	607	S -> L (IN REF. 1).
FT	CONFLICT	691	691	R -> H (IN REF. 1).
50	SEQUENCE	695 AA;	78172 NM;	ESBEBDB29C70C450 CRC64;

	Query Marc3	4.4%	Score 104.5;	DB 1;	Length 655;
	Best Local Similarity	16.5%	Pred No. 1.9;	Mismatches 169;	Indels 199; Gaps 21
	Matches	91;	Conservative	91;	
QY	18 LPNG-INGIKDAKVTGVIGSGDFPAKSLTTLIRICGHHVYIGSRNKPFASEFPFPHVDV	76			
Db	61 IPGASFGSGDDEKI-----EISQNDVEIEAN-----VFENLPEKL-----	97			
QY	77 THHEDALTKNLIIIVAIHREHTYSIMDRHLVG-----KILIDSVNN	119			
Db	98 --HEIIEIRANNL-LVIDPDAFOINPNRYLLINSNTGVKRLPAVKIKQSLQXKVLADIQON	154			
QY	120 MRINQRPESNAEYLSTLPFDLSLVIGFVVAVAMALQLGPDKASROYIGENNIOARQCVI	179			
Db	155 INIH-----TVERNSPVGLSFPSMLTM-----SKNGIHEHNCA-----	189			
QY	180 ELARKQNFPIPIDGISSSAREIENIPLEFLFMGEFVVVAISLATFFLYSFTVDVIHPY	239			
Db	190 ----ENGQLOELINSDNDNIETPLPDNVDFQAGSGPVILLISRTRISHSPYLENIMKTL	244			
QY	240 ARRQQSDPFKIP-----EIVNKTL-----PIVATITL-	267			
Db	245 RAKSTYNLKLTPELEKFVTLMASTLYSHCCAPANMRROISDLHPICNSIIRGEVDVM	304			

QY	268	-----SLVYLLAGLAAAYQLYGTGKXRF-----	-----PPLTEFLQCKXOLG-----	300
Db	305	TOANGQRSLABDDGSSILAGSPDTMYSPFDLDCNEVDVICSPEPDITFNPCEDDMGHDI		364
QY	307	--LISFFPAMVVA-----	-----YSLCLPMRBSERYFLNMAVQOVHANINSMNE	351
Db	365	LRVLWIFSLAIGNIIVILLITRSQXLTVF-----	-----RFLMCNLAFAAD-----	408
QY	352	EEVWRIEWYISFGMSIGLISLIVATSI.PSVSNALNMR-----	-----FSFIQSLTG----	399
Db	409	-----LIGIYLLILIASVDHTKTYQHNVAIDWQTGAGCDAAAGFTWFASELSVYT		459
QY	400	YVALLISTFH-----	-----VLIVGKRAFEERYFRYPPEVFLATVLP	439
Db	460	LTAATLBRWHTITHAMQLOCKVKQVRHASMVLVGMIFAFVYALPFLFGISSVMKXISICLP		519
QY	440	STVLIDLQL 449		
Db	520	MDISPLSQL 529		
RESULT 11				
SOTB_ECO57				
ID	_SOTB_ECO57	STANDARD;	PRT;	396 AA.
AC	P58529;			
DT	28-FEB-2003 (Rel. 41, Created)			
DT	28-FEB-2003 (Rel. 41, Last sequence update)			
DT	15-MAR-2004 (Rel. 43, Last annotation update)			
DE	Sugar efflux transporter.			
GN	SOTB OR Z2173 OR ECG2135.			
OS	Escherichia coli O157:H7.			
OC	Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;			
OC	Enterobacteriaceae; Escherichia.			
OX	NCBI_TaxID=83334;			
EN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=O157:H7 / EDL933 / ATCC 700927;			
EX	MEDLINE=21074935; PubMed=11206551;			
RA	Perna N.T., Plunkett G. III, Burland V., Mau B., Glaesner J.D.,			
RA	Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,			
RA	Posefai G., Hackett J., Klink S., Boutin A., Shao Y., Miller L.,			
RA	Grotbeck E.J., Davis N.W., Lim A., Dimalanta E.T., Potamouzis K.,			
RA	Apodaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C.,			
RA	Weich R.A., Blattner F.R.;			
RT	"Genome sequence of enterohaemorrhagic Escherichia coli O157:H7.";			
RL	Nature 409:529-533(2001).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=O157:H7 / RIMD 0509952;			
EX	MEDLINE=21156231; PubMed=11258796;			
RA	Hayashi T., Makino K., Omishi M., Kurokawa K., Ishii K., Yokoyama K.,			
RA	Han C.-G., Ohtsudo E., Nakayama K., Murata T., Tanaka M., Tobe T.,			
RA	Ikida T., Takami H., Honda T., Sasekawa C., Ogasawara N., Yasunaga T.,			
RA	Kumata S., Shibata T., Hattori M., Shinagawa H.;			
RT	"Complete genome sequence of enterohaemorrhagic Escherichia coli			
RT	O157:H7 and genomic comparison with a laboratory strain K-12.";			
RL	DNA Res. 8:11-22(2001).			
CC	-1- FUNCTION: Involved in the efflux of sugars. The physiological role			
CC	may be the reduction of the intracellular concentration of toxic			
CC	sugars or sugar metabolites. Transports L-arabinose and to a			
CC	lesser extent IPTG. Seems to contribute to the control of the			
CC	arabinose regulon (by similarity).			
CC	-1- SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane			
CC	(probable).			
CC	-1- SIMILARITY: Belongs to major facilitator superfamily. SotB			
CC	(TC 2.A.1.2) family.			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -			
CC	the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement (See <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a>			

```

CC or send an email to license@lsb-sib.ch).
CC -----
DR EMBL; AE005552; AAC56236.1; -
DR EMBL; AF002557; BAB35558.1; -
DR PIR; G90895; G90895.
DR PIR; H85721; H85721.
DR HAMAP; MF 00517; -; 1.
DR InterPro; IPR007114; MFS.
DR InterPro; IPR005828; Sub_transporter.
DR Pfam; PF00083; Sugar_Ltr; 1.
DR PROSITE; PS50850; MFS; 1.
KM Transport; Sugar transport; Transmembrane; Inner membrane;
KM Complete proteome.
FT TRANSMEM 15 35 POTENTIAL.
FT TRANSMEM 50 70 POTENTIAL.
FT TRANSMEM 81 101 POTENTIAL.
FT TRANSMEM 103 123 POTENTIAL.
FT TRANSMEM 136 156 POTENTIAL.
FT TRANSMEM 170 190 POTENTIAL.
FT TRANSMEM 209 229 POTENTIAL.
FT TRANSMEM 246 266 POTENTIAL.
FT TRANSMEM 275 295 POTENTIAL.
FT TRANSMEM 319 339 POTENTIAL.
FT TRANSMEM 333 353 POTENTIAL.
FT TRANSMEM 364 384 POTENTIAL.
SQ SEQUENCE 396 AA; 42568 MW; CB6A209B0BE6D4F0 CRC64;

Query Match 4.4%; Score 104; DB 1; Length 396;
Best Local Similarity 18.8%; Pred. No. 1;
Matches 85; Conservative 72; Mismatches 148; Indels 146; Gaps 20;

QY 31 VTVVGIGSDGPAKSLTIRLRGCVHVVIGSRNPKASEFPFVVDVTHEDALTKTNIF 90
Db 31 VPGGLT--SDIAQSFHQTAQVGMILTIVAMVVALMSLPFMTSGVERBKILICLFVVF 88

QY 91 VAHREHYTSMDRLVCKILLIDVSNMKNQVPEBAEFLASLPDSLIYKFPNVVS 150
Db 89 LASHVLSFLS-WSFTVLVISRI-----GVAPAHAI-----WSITA 123

QY 151 AVALQUGP--KQASROVYICSNNIQARQVIELAR-----QNFIPIDIGSL----- 195
Db 124 SLAIRAPAKKRAQALSLATGALAMVGLGRIVGQFGKMTFFAIGALITLLC 183

QY 196 -----SSAREENLPRLFTMGQPVVAISLAT-----FFPLYSFVDVHPY 239
Db 184 LKLLPLPSEHSGSLKSLPL-----LFRBPALMSIYLLTVVVVTAHYAYSY-----IIPF 235

QY 240 AER---QOSDFYKIPLEIVNKLPIVAITLLSLVYLAGLLAAAYQLYGTXYRRPML 296
Db 236 VQNIAGFSANF-----ATLMLLLGGAGITGIVTGKLGNOYAS----- 274

QY 297 TWLQCRKQLGLSFPFAMVAVSLCLPMSRSERYLFLMAYQVAVANIENSMNEEYVR 356
Db 275 -----ALVSTALALLVCLALLPANSE-----IHLGVLISF-----WG 309

QY 357 IENYISFGIMSLGLSL-----LAVTSIPVSN-----ALNREBSFIQST 397
Db 310 IAMMILGLQKQKVALPADATDVAAALFSGIFNIGIGAGALVENQVSLHWS-----MSM 364

QY 398 LGYVALLISTFVLIYV-----WKRAPEE 422
Db 365 IGYVG-TVPAFALIMSIIFPRMPVTLDEQ 394

RESULT 12
SOTB_ECOLI STANDARD; PRT; 396 AA.
AC P31122; P76883; P77353;
DT 01-JUL-1993 (Rel. 26, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Sugar efflux transporter.
GN SOTB OR B1528.

```

```

OC Escherichia coli.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=K12 / MG1655;
RX MEDLINE=97426617; PubMed=9276503;
RA Blatter F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
RA Mau B., Shao Y.;
RT "The complete genome sequence of Escherichia coli K-12.";
RL Science 277:1453-1474(1997).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=K12;
RX MEDLINE=97251357; PubMed=9097039;
RA Alba H., Baba T., Fujita K., Hayashi K., Inada T., Isono K., Itoh T.,
RA Kaeai H., Kashimoto K., Kimura S., Kitakawa M., Kitagawa M.,
RA Nakano K., Miki T., Mizobuchi K., Mori H., Mori T., Motomura K.,
RA Nakade S., Nakamura Y., Nishimoto H., Nishio Y., Oshima T., Saito N.,
RA Sempel G., Seki Y., Sivastandaram S., Tagami H., Takeda J.,
RA Takemoto K., Takeuchi Y., Wada C., Yamamoto Y., Horiuchi T.;
RA "A 570-kb DNA sequence of the Escherichia coli K-12 genome
corresponding to the 28.0-40.1 min region on the linkage map.";
RL DNA Res. 3:363-377(1996).
RN [3]
RP SEQUENCE OF 217-396 FROM N.A.
RX MEDLINE=93186717; PubMed=8383113;
RA Cohen S.P., Haechler H., Levy S.B.;
RT "Genetic and functional analysis of the multiple antibiotic
resistance (mar) locus in Escherichia coli.";
RL J. Bacteriol. 175:1484-1492(1993).
RN [4]
RP CHARACTERIZATION.
RC STRAIN=SB0;
RX MEDLINE=99194728; PubMed=10094697;
RA Boet S., Silva F., Belin D.;
RT "Transcriptional activation of ydeA, which encodes a member of the
major facilitator superfamily, interferes with arabinose accumulation
and induction of the Escherichia coli arabinose PBAD promoter.";
RL J. Bacteriol. 181:2185-2191(1999).
RN [5]
RP CHARACTERIZATION.
RC STRAIN=J8219;
RX MEDLINE=99369894; PubMed=10438792;
RA Carole S., Pichoff S., Bouche J.-P.;
RT "Escherichia coli gene ydeA encodes a major facilitator pump which
exports L-arabinose and isopropyl-beta-D-thiogalactopyranoside.";
RL J. Bacteriol. 181:5123-5125(1999).
CC -1- FUNCTION: Involved in the efflux of sugars. The physiological role
may be the reduction of the intracellular concentration of toxic
sugars or sugar metabolites. Transports L-arabinose and to a
lesser extent IPTG. Seems to contribute to the control of the
arabinose regulon.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane
(probable).
CC -1- SIMILARITY: Belongs to major facilitator superfamily. SotB
(TC 2.A.1.2) family.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation-
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@lsb-sib.ch).
CC -----
DR EMBL; AE000250; AAC74601.1; -
DR EMBL; D90795; BAA15210.1; -
DR EMBL; D90796; BAA15218.1; -
DR EMBL; D90797; BAA15230.1; -

```

```

DR EMBL: M96235; NOT_ANNOTATED_CDS.
DR PIR: C64907; C64907.
DR EcoGene: EG11636; socB.
DR HAMAP: MF_00517; 1.
DR InterPro: IPR007114; MFS.
DR InterPro: IPR005828; Sub transporter.
DR Pfam: PF00083; sugar tr_1.
DR PROSITE: PS00650; MFS_1.
KM Transport; Sugar transporter; Transmembrane; Inner membrane;
KM Complete proteome.
FT TRANSMEM 15 35 POTENTIAL.
FT TRANSMEM 50 70 POTENTIAL.
FT TRANSMEM 81 101 POTENTIAL.
FT TRANSMEM 103 123 POTENTIAL.
FT TRANSMEM 136 156 POTENTIAL.
FT TRANSMEM 170 190 POTENTIAL.
FT TRANSMEM 209 229 POTENTIAL.
FT TRANSMEM 246 266 POTENTIAL.
FT TRANSMEM 275 295 POTENTIAL.
FT TRANSMEM 299 319 POTENTIAL.
FT TRANSMEM 333 353 POTENTIAL.
FT TRANSMEM 364 384 POTENTIAL.
SQ SEQUENCE 396 AA; 42538 MW; CB6A34CA4E6D4F0 CRC64;

Query Match 4.4%; Score 104; DB 1; Length 396;
Best Local Similarity 18.8%; Pred. No. 1;
Matches 85; Conservative 72; Mismatches 148; Indels 146; Gaps 20;

QY 31 VTGVIIGSDPAKSLTIRLCRGYHVVIGSRNPKASEFPFVVDVTHEBALTKNIF 90
DB 31 VVGGLL--SDIAQSFMOQAQVIGMLTYAVVALMSLPFMLMTQVERRKILICLFVF 88
QY 91 VAHREHYTSLMDLHLVGLKILIDVSNMRINQVPESSAEYLAFLPDSLSLVGENVVS 150
DB 89 IASHVLSPUS-WSFYVLVLSRI-----GVAFNAHIF-----WSITA 123
QY 151 AVALDQGP--KDAEQVYICSNMICAQOVLRLAR-----QINFTPIDLGSF---- 195
DB 124 SLAIMAPAKKAQAQLSLATGTALAVGLPLGIVQVGMWTFPAIGALITLLC 183
QY 196 -----SARIEINPLRLFTLMRGVVAISLAT-----FFLYSPVDVTHPY 239
DB 184 LIKLPLPSEHSGSLKSLPL--LFRFPALMS-YLLTVVVVTAHYAIVSY----IEPF 235
QY 240 AEN--COQDFYKPIEIVNKTLPVATTLISLVYLAGLAAVQLYGTGKRRFPWLE 296
DB 236 VQNIQFSGANF-----ACALLLLGGAGIIGSVIRGKLGNYAS----- 274
QY 297 TWLQCRKQGLISFPFAMVHVAISCLPFRSERYLFLNMAVQOVHANINSEWEEVWR 356
DB 275 -----ALVSTAILLVCLALLLPANSE-----IHLGVLSIF-----WG 309
QY 357 IEMYISFGIMSLGLSL-----LAVTSIPSVSN-----ALMNRSESFIOST 397
DB 310 IAMMIGMGQYKVALAPADATDVAMALFSGIFNIGTAGALVGNQVSLHWS-----MSM 364
QY 398 LGYVALLISTFHVLIYG-----WRAPEEE 422
DB 365 IGYVG-AVPAFAALIMSIIFRRMPVLTLEQ 394

RESULT 13
NPTL_MOUSE STANDARD; PRT; 465 AA.
AC Q61983;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Renal sodium-dependent phosphate transporter protein 1 (Sodium/phosphate cotransporter 1) (Na(+)/Pi cotransporter 1) (Renal sodium-phosphate cotransporter 1) (Renal Na(+)-dependent phosphate cotransporter 1).
DE SLCT17A1 OR NPTL.
GN

```

```

OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_Taxid=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Kidney;
RA MEDLINE=95353846; PubMed=7611445;
RX Chong S.S., Kozak C.A., Liu L., Kristjansson K., Dunn S.T.,
RA Bourdeau J.E., Hughes M.R.;
RT "Cloning, genetic mapping, and expression analysis of a mouse renal
RL sodium-dependent phosphate cotransporter."
CC Am. J. Physiol. 268:F1038-F1045(1995).
CC -1- FUNCTION: Important for the resorption of phosphate by the kidney.
CC May be involved in actively transporting phosphate into cells via
CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
CC -1- TISSUE SPECIFICITY: Kidney.
CC -1- SIMILARITY: Belongs to the sodium/phosphate cotransporter family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X71241; CA54459.1; -.
DR PIR: S69915; S69915.
DR MGD: MGI:103209; Slc17a1.
DR InterPro: IPR007114; MFS.
DR InterPro: IPR004745; Pi cotransp.
DR TIGRFAMs: TIGR00894; 2A014euk; 1.
DR PROSITE: PS00850; MFS_1.
KM Transport; Symport; Sodium transporter; Transmembrane; Glycoprotein.
FT TRANSMEM 79 99 POTENTIAL.
FT TRANSMEM 117 137 POTENTIAL.
FT TRANSMEM 176 196 POTENTIAL.
FT TRANSMEM 199 219 POTENTIAL.
FT TRANSMEM 260 280 POTENTIAL.
FT TRANSMEM 304 324 POTENTIAL.
FT TRANSMEM 337 356 POTENTIAL.
FT TRANSMEM 363 383 POTENTIAL.
FT TRANSMEM 399 419 POTENTIAL.
FT TRANSMEM 429 449 POTENTIAL.
FT CARBOHYD 39 39 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 47 47 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 56 56 N-LINKED (GLCNAC. .) (POTENTIAL).
SQ SEQUENCE 465 AA; 51589 MW; C67BE25A2C291EEF CRC64;

Query Match 4.4%; Score 103; DB 1; Length 465;
Best Local Similarity 18.7%; Pred. No. 1.5; Indels 124; Gaps 16;
Matches 76; Conservative 64; Mismatches 142;

QY 110 GKILIDVSNMRINQVPESSAEYLAFLPDSLSLVGENVVSAMALQGPXA---SRQVY 166
DB 79 GLITSVFPGMYVQAP---VGLISGIYPMKILIGSLSLISLMLIPPAQVAAVYI 135
QY 167 ICSNNTQARQVIELARQNF-----PIDGLSSAKREIENPLRLFTLMRGVVA---- 218
DB 136 VCRVLIQIAQGTVSTGQHEIWKMAPLERGLTS-----VTLGQFVN--GPFYVLVS 187
QY 219 -----AISTATFFFLTSFVADVTHVYPARNQOSDF----- 247
DB 188 GFICDLGMPWFIYFGIVGCVLSLSMFLFPDDPRD--HPMSSSEKXIISSLMQAS 245
QY 248 ---YKPIPIVANKTLPVATILS----- 268
DB 246 SGRQSLPIYAKMKSPLMIIINSFAFISNSLVTYPTFTISTVLAHVAVRANGLSLSP 305
QY 269 --LVYLAGLAAVQLYGTGKRRFPFPMLETLQCRKQGLISFPFAMVHVAISCLPWR 326

```

QY 186 NFIPIDGLSSAREIENLPRLFTLWGRGPVVAISLATFFFLYSFVRDVIHPYARNOOS 245

21	1KANSWEM	1/9	198	POTENTIAL.
50	SEQUENCE	220 AA:	24815 MW:	182C0244743B1

Query Match 4.3%; Score 102; DB 1; Length 220;  
Best Local Similarity 23.8%; Pred. No. 0.72;  
Matches 43; Conservative 25; Mismatches 51; Indels 62; Gaps 7;

QY	267	LSLVYLAGLAAAYOLYGT-----KRRPPMLETW-----	298
Db	20	LWLVTAGFVPVWTFYLGATGQLGADPVKTFEHLGMLRFLITLVTMRDLGIT	79
QY	299	-LQCKKQGLSFFPAMVAVSLCPMRSERYFLNNAYQOVHANIENSNNEEYWR	357
Db	80	LRRYRAGLLAFYALMHT-----TWVLDGMLLSAITDIYR-	120
QY	358	EWYISFGIMSLGISLAVTSIPSYNLMW-----RPSFIQSTLGYVALISTFHYL	411
Db	121	RPFITGMISLALIVPLALTS-----NMSIRKLGRRWSLHK-LVYIAAGSAVHFL	172
QY	412	I 412	
Db	173	M 173	

Search completed: March 1, 2004, 23:31:14  
Job time : 44 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 1, 2004, 23:34:47 ; Search time 63 Seconds

(without alignments)  
1521.644 Million cell updates/sec

Title: US-09-455-486-6

Perfect score: 2351

Sequence: 1 MESISMGPSKSLSETCLPN.....ALVPSIVLDLQICRPD 454

Scoring table: BLOSUM62

Gapop 10.0, Gapext 0.5

Searched: 809742 seqs, 21153259 residues

Total number of hits satisfying chosen parameters: 809742

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%  
Listing first 45 summaries

Database :

Published Applications AA:\*  
1: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB.pep:\*  
2: /cgn2\_6/ptodata/2/pubpaa/PCT\_NEW\_PUB.pep:\*  
3: /cgn2\_6/ptodata/2/pubpaa/US06\_NEW\_PUB.pep:\*  
4: /cgn2\_6/ptodata/2/pubpaa/US06\_PUBCOMB.pep:\*  
5: /cgn2\_6/ptodata/2/pubpaa/PCTUS\_NEW\_PUB.pep:\*  
6: /cgn2\_6/ptodata/2/pubpaa/PCTUS\_PUBCOMB.pep:\*  
7: /cgn2\_6/ptodata/2/pubpaa/US08\_NEW\_PUB.pep:\*  
8: /cgn2\_6/ptodata/2/pubpaa/US08\_PUBCOMB.pep:\*  
9: /cgn2\_6/ptodata/2/pubpaa/US09\_PUBCOMB.pep:\*  
10: /cgn2\_6/ptodata/2/pubpaa/US09\_PUBCOMB.pep:\*  
11: /cgn2\_6/ptodata/2/pubpaa/US09C\_PUBCOMB.pep:\*  
12: /cgn2\_6/ptodata/2/pubpaa/US09\_NEW\_PUB.pep:\*  
13: /cgn2\_6/ptodata/2/pubpaa/US10\_PUBCOMB.pep:\*  
14: /cgn2\_6/ptodata/2/pubpaa/US10C\_PUBCOMB.pep:\*  
15: /cgn2\_6/ptodata/2/pubpaa/US10C\_PUBCOMB.pep:\*  
16: /cgn2\_6/ptodata/2/pubpaa/US10C\_NEW\_PUB.pep:\*  
17: /cgn2\_6/ptodata/2/pubpaa/US60\_NEW\_PUB.pep:\*  
18: /cgn2\_6/ptodata/2/pubpaa/US60\_PUBCOMB.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match Length	ID	Description
1	2351	100.0	454 10 US-09-888-2574-10	Sequence 10, Appl
2	2351	100.0	454 14 US-10-165-044-8	Sequence 8, Appl
3	2351	100.0	454 15 US-10-239-607-29	Sequence 29, Appl
4	2351	100.0	454 15 US-10-295-027-574	Sequence 574, Appl
5	2351	100.0	454 15 US-10-382-190-2	Sequence 2, Appl
6	2348	99.9	454 15 US-10-104-047-3155	Sequence 3155, Ap
7	2294	97.6	490 9 US-09-802-520-1	Sequence 1, Appl
8	2290	97.4	490 15 US-10-239-607-14	Sequence 14, Appl
9	2036	86.6	419 15 US-10-239-607-32	Sequence 32, Appl
10	1272	54.1	488 14 US-10-182-951-6	Sequence 6, Appl
11	1269	54.0	488 9 US-09-747-835A-13	Sequence 13, Appl
12	1269	54.0	488 15 US-10-239-607-36	Sequence 36, Appl
13	1269	54.0	488 15 US-10-239-607-38	Sequence 38, Appl
14	1088	46.3	459 9 US-09-965-529-21	Sequence 21, Appl
15	1088	46.3	459 10 US-09-969-680A-21	Sequence 21, Appl

16	1087	46.2	459 14 US-10-165-044-10	Sequence 10, Appl
17	1073	45.6	461 15 US-10-239-607-34	Sequence 34, Appl
18	901	38.3	173 14 US-10-011-095-8	Sequence 8, Appl
19	901	38.3	173 14 US-10-010-667A-8	Sequence 8, Appl
20	846	36.0	375 10 US-09-895-298-162	Sequence 162, Appl
21	844	35.9	179 9 US-09-864-761-37654	Sequence 37654, A
22	766	32.7	237 9 US-09-747-835A-15	Sequence 15, Appl
23	736	31.3	111 9 US-09-363-896-1	Sequence 1, Appl
24	717	30.5	267 9 US-09-747-835A-50	Sequence 50, Appl
25	717	30.5	267 9 US-09-747-835A-51	Sequence 51, Appl
26	717	30.5	339 9 US-09-759-143-879	Sequence 879, Appl
27	717	30.5	339 9 US-09-780-669-879	Sequence 879, Appl
28	717	30.5	339 9 US-09-822-827-879	Sequence 879, Appl
29	717	30.5	339 9 US-09-802-520-11	Sequence 11, Appl
30	717	30.5	339 9 US-09-895-793-879	Sequence 879, Appl
31	717	30.5	339 9 US-09-895-814-879	Sequence 879, Appl
32	717	30.5	339 13 US-10-012-896-879	Sequence 879, Appl
33	717	30.5	339 14 US-10-011-095-2	Sequence 2, Appl
34	717	30.5	339 14 US-10-010-667A-2	Sequence 2, Appl
35	717	30.5	339 14 US-10-205-823-397	Sequence 397, Appl
36	717	30.5	339 14 US-10-144-678A-879	Sequence 879, Appl
37	717	30.5	339 14 US-10-294-025-879	Sequence 879, Appl
38	717	30.5	339 15 US-10-239-607-37	Sequence 37, Appl
39	717	30.5	339 15 US-10-295-027-714	Sequence 714, Appl
40	717	30.5	339 15 US-10-295-027-1347	Sequence 1347, Ap
41	717	30.5	375 14 US-10-165-044-2	Sequence 2, Appl
42	628	26.7	283 14 US-10-043-487-313	Sequence 313, Appl
43	628	26.7	283 14 US-10-205-823-123	Sequence 123, Appl
44	628	26.7	283 15 US-10-239-607-39	Sequence 39, Appl
45	433	18.4	232 10 US-09-895-298-80	Sequence 80, Appl

#### ALIGNMENTS

RESULT 1  
US-09-888-2574-10  
Sequence 10, Application US/098882574  
; Application No. US20030060612A1  
GENERAL INFORMATION:  
APPLICANT: Goddard, Audrey  
APPLICANT: Godowski, Paul J.  
APPLICANT: Gurney, Austin L.  
APPLICANT: Hillan, Kenneth J.  
APPLICANT: Polakis, Paul  
APPLICANT: Smith, Victoria  
APPLICANT: Wood, William I.  
APPLICANT: Wu, Thomas D.  
APPLICANT: Zhang, Zemin  
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DIAGNOSIS AND  
FILE REFERENCE: P5002R1  
CURRENT APPLICATION NUMBER: US/09/888,2574  
CURRENT FILING DATE: 2002-04-30  
PRIOR APPLICATION NUMBER: US 60/063,540  
PRIOR FILING DATE: 1997-10-28  
PRIOR APPLICATION NUMBER: US 60/089,653  
PRIOR FILING DATE: 1998-06-17  
PRIOR APPLICATION NUMBER: US 60/099,792  
PRIOR FILING DATE: 1998-09-10  
PRIOR APPLICATION NUMBER: US 60/103,678  
PRIOR FILING DATE: 1998-10-08  
PRIOR APPLICATION NUMBER: US 60/235,451  
PRIOR FILING DATE: 2000-09-26  
PRIOR APPLICATION NUMBER: PCT/US99/12252  
PRIOR FILING DATE: 1999-06-02  
PRIOR APPLICATION NUMBER: PCT/US99/20111  
PRIOR FILING DATE: 1999-09-01  
PRIOR APPLICATION NUMBER: PCT/US00/04342  
PRIOR FILING DATE: 2000-02-18  
PRIOR APPLICATION NUMBER: PCT/US00/05841  
PRIOR FILING DATE: 2000-03-02  
PRIOR APPLICATION NUMBER: PCT/US00/08439

NOT PUBLISHED

disclosed here



PRIOR FILING DATE: 2000-03-30  
PRIOR APPLICATION NUMBER: PCT/US00/23328  
PRIOR FILING DATE: 2000-08-24  
PRIOR APPLICATION NUMBER: PCT/US00/32678  
PRIOR FILING DATE: 2000-12-01  
PRIOR APPLICATION NUMBER: PCT/US01/06520  
PRIOR FILING DATE: 2001-02-28  
PRIOR APPLICATION NUMBER: PCT/US01/06666  
PRIOR FILING DATE: 2001-03-01  
NUMBER OF SEQ ID NOS: 10  
SEQ ID NO 10  
LENGTH: 454  
TYPE: PRT  
ORGANISM: Homo Sapien  
US-09-888-257A-10

Query Match 100.0%; Score 2351; DB 10; Length 454;  
Best Local Similarity 100.0%; Pred. No. 8e-222;  
Matches 454; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MESISMWGPXKLSSETCLPENGINGIDAKRVTVGVIGSGDPKASLTIRIRCGYHVIGS 60  
DB 1 MESISMWGPXKLSSETCLPENGINGIDAKRVTVGVIGSGDPKASLTIRIRCGYHVIGS 60  
QY 61 RNPKEASEFPFPHVVDVTHEDALTKNIIFFVAIHREHYTSLMDRLHVLGKILIDVSNM 120  
DB 61 RNPKEASEFPFPHVVDVTHEDALTKNIIFFVAIHREHYTSLMDRLHVLGKILIDVSNM 120  
QY 121 RINQYPSNAEYIASLPDLSLYKGFNVYSAMALQGPCKASQVYICSNNIQARQVIE 180  
DB 121 RINQYPSNAEYIASLPDLSLYKGFNVYSAMALQGPCKASQVYICSNNIQARQVIE 180  
QY 181 LARQNLPIPDIGLSASAREIENLPLRLFTLWGPVVAISLATPEFLYSFVADVHPYA 240  
DB 181 LARQNLPIPDIGLSASAREIENLPLRLFTLWGPVVAISLATPEFLYSFVADVHPYA 240  
QY 241 RNOQSDPYKPIPIEVNKTLPYVATITLSLYLAGLAAAYQYGYKRRPFWLETLQ 300  
DB 241 RNOQSDPYKPIPIEVNKTLPYVATITLSLYLAGLAAAYQYGYKRRPFWLETLQ 300  
QY 301 CRKQGLISFFPMVHVAAYSLCLPMRSEHYLFLNAAQOVYHANIENSNMEEVWRIEM 360  
DB 301 CRKQGLISFFPMVHVAAYSLCLPMRSEHYLFLNAAQOVYHANIENSNMEEVWRIEM 360  
QY 361 ISFGIMSLGLSLAATSIPIVSNAIWMREFSFIQSTLGYVALLISTFHVLIYGMKRAFE 420  
DB 361 ISFGIMSLGLSLAATSIPIVSNAIWMREFSFIQSTLGYVALLISTFHVLIYGMKRAFE 420  
QY 421 EBYRFFYTPNPFVLAIVLPSIVILDLQLCRYPD 454  
DB 421 EBYRFFYTPNPFVLAIVLPSIVILDLQLCRYPD 454

RESULT 2  
US-10-165-044-8  
Sequence 8, Application US/10165044  
Publication No. US20030149531A1  
GENERAL INFORMATION:  
APPLICANT: Agensys, Inc.  
APPLICANT: Rene S. Hubert  
APPLICANT: Arthur B. Raitano  
APPLICANT: Douglas Saffran  
APPLICANT: Daniel E.H. Afar  
APPLICANT: Steven Chappell Mitchell  
APPLICANT: Mary Paris  
APPLICANT: Nya Jakobovits  
TITLE OF INVENTION: NOVEL SERPENTINE TRANSMEMBRANE ANTIGENS  
FILE REFERENCE: 51158-20016.02  
CURRENT APPLICATION NUMBER: US/10/165,044  
PRIOR FILING DATE: 2002-06-06  
PRIOR APPLICATION NUMBER: US 60/087,520  
PRIOR FILING DATE: 1998-06-01

PRIOR APPLICATION NUMBER: US 60/091,183  
PRIOR FILING DATE: 1998-06-30  
PRIOR APPLICATION NUMBER: US 09/323,873  
PRIOR FILING DATE: 1999-06-01  
PRIOR APPLICATION NUMBER: US 09/455,486  
PRIOR FILING DATE: 1999-12-06  
PRIOR APPLICATION NUMBER: WO 99/62941  
PRIOR FILING DATE: 1999-06-01  
PRIOR APPLICATION NUMBER: PCT/US00/33040  
PRIOR FILING DATE: 2000-12-06  
NUMBER OF SEQ ID NOS: 57  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 8  
LENGTH: 454  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-10-165-044-8

Query Match 100.0%; Score 2351; DB 14; Length 454;  
Best Local Similarity 100.0%; Pred. No. 8e-222;  
Matches 454; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MESISMWGPXKLSSETCLPENGINGIDAKRVTVGVIGSGDPKASLTIRIRCGYHVIGS 60  
DB 1 MESISMWGPXKLSSETCLPENGINGIDAKRVTVGVIGSGDPKASLTIRIRCGYHVIGS 60  
QY 61 RNPKEASEFPFPHVVDVTHEDALTKNIIFFVAIHREHYTSLMDRLHVLGKILIDVSNM 120  
DB 61 RNPKEASEFPFPHVVDVTHEDALTKNIIFFVAIHREHYTSLMDRLHVLGKILIDVSNM 120  
QY 121 RINQYPSNAEYIASLPDLSLYKGFNVYSAMALQGPCKASQVYICSNNIQARQVIE 180  
DB 121 RINQYPSNAEYIASLPDLSLYKGFNVYSAMALQGPCKASQVYICSNNIQARQVIE 180  
QY 181 LARQNLPIPDIGLSASAREIENLPLRLFTLWGPVVAISLATPEFLYSFVADVHPYA 240  
DB 181 LARQNLPIPDIGLSASAREIENLPLRLFTLWGPVVAISLATPEFLYSFVADVHPYA 240  
QY 241 RNOQSDPYKPIPIEVNKTLPYVATITLSLYLAGLAAAYQYGYKRRPFWLETLQ 300  
DB 241 RNOQSDPYKPIPIEVNKTLPYVATITLSLYLAGLAAAYQYGYKRRPFWLETLQ 300  
QY 301 CRKQGLISFFPMVHVAAYSLCLPMRSEHYLFLNAAQOVYHANIENSNMEEVWRIEM 360  
DB 301 CRKQGLISFFPMVHVAAYSLCLPMRSEHYLFLNAAQOVYHANIENSNMEEVWRIEM 360  
QY 361 ISFGIMSLGLSLAATSIPIVSNAIWMREFSFIQSTLGYVALLISTFHVLIYGMKRAFE 420  
DB 361 ISFGIMSLGLSLAATSIPIVSNAIWMREFSFIQSTLGYVALLISTFHVLIYGMKRAFE 420  
QY 421 EBYRFFYTPNPFVLAIVLPSIVILDLQLCRYPD 454  
DB 421 EBYRFFYTPNPFVLAIVLPSIVILDLQLCRYPD 454

RESULT 3  
US-10-239-607-29  
Sequence 29, Application US/10239607  
Publication No. US20030219761A1  
GENERAL INFORMATION:  
APPLICANT: Saatchiglu, Fahri  
TITLE OF INVENTION: No. US20030219761A1 Prostate-Specific or  
TITLE OF INVENTION: Testis-Specific Nucleic Acid Molecules, Polypeptides, and  
FILE REFERENCE: 50218/003002  
CURRENT APPLICATION NUMBER: US/10/239,607  
PRIOR FILING DATE: 2002-09-24  
PRIOR APPLICATION NUMBER: PCT/US01/09410  
PRIOR FILING DATE: 2001-03-23  
PRIOR APPLICATION NUMBER: US 60/191,929  
PRIOR FILING DATE: 2000-03-24  
NUMBER OF SEQ ID NOS: 73  
SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 29  
LENGTH: 454  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-10-239-607-29

Query Match 100.0%; Score 2351; DB 15; Length 454;  
Best Local Similarity 100.0%; Pred. No. 8e-222;  
Matches 454; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 1 MESISMWSPKSLSETCLPENGINGIDARKVTGVIGSGDPFASLTIRLRGCVHYVIGS 60
DB 1 MESISMWSPKSLSETCLPENGINGIDARKVTGVIGSGDPFASLTIRLRGCVHYVIGS 60
QY 61 RNPKFASFPFHVDYTHEDALTKTNIIFVAIHREHYTSLMDRLRLVKGKILIDVSNM 120
DB 61 RNPKFASFPFHVDYTHEDALTKTNIIFVAIHREHYTSLMDRLRLVKGKILIDVSNM 120
QY 121 RINQVPSNAEYIASLFPDLSLYKGFNVSAWALQGPDAQRQVYICSNNOARQOYIE 180
DB 121 RINQVPSNAEYIASLFPDLSLYKGFNVSAWALQGPDAQRQVYICSNNOARQOYIE 180
QY 181 LARQNFIPIDGSLSSAREIENLPLRLFTLMRGVVAVVAISLATEFFLYSFVADVHPYA 240
DB 181 LARQNFIPIDGSLSSAREIENLPLRLFTLMRGVVAVVAISLATEFFLYSFVADVHPYA 240
QY 241 RNOQSDFYKIPFIEVNTKLPYVAITLLSLVYLAGLAAAYQLYGKRRPFWLSTWLO 300
DB 241 RNOQSDFYKIPFIEVNTKLPYVAITLLSLVYLAGLAAAYQLYGKRRPFWLSTWLO 300
QY 301 CRKQGLISFPFAMVAVASLCLPMRSEERYFLNMAVQOVANIENTSNWEEVWRIEMY 360
DB 301 CRKQGLISFPFAMVAVASLCLPMRSEERYFLNMAVQOVANIENTSNWEEVWRIEMY 360
QY 361 ISFGIMSLGSLAVTSIPSVSNALNMBREPSFIQSTLGVALISTFHVLIYGMKRAFE 420
DB 361 ISFGIMSLGSLAVTSIPSVSNALNMBREPSFIQSTLGVALISTFHVLIYGMKRAFE 420
QY 421 EBYRFFYTPNPFVLAIVLPSIYVLDLQCRYPD 454
DB 421 EBYRFFYTPNPFVLAIVLPSIYVLDLQCRYPD 454
```

RESULT 4  
US-10-295-027-574  
Sequence 574, Application US/10295027  
Publication No. US2003023250A1

GENERAL INFORMATION:

APPLICANT: Afar, Daniel  
APPLICANT: Aziz, Natacha  
APPLICANT: Ginsberg, Wendy M.  
APPLICANT: Gish, Kurt C.  
APPLICANT: Glynn, Richard  
APPLICANT: Hevezi, Peter A.  
APPLICANT: Mack, David H.  
APPLICANT: Murray, Richard  
APPLICANT: Watson, Susan R.  
APPLICANT: Eos Biotechnology, Inc.  
TITLE OF INVENTION: Methods of Diagnosis of Cancer, Compositions and  
FILE REFERENCE: 018501-012500US  
CURRENT APPLICATION NUMBER: US/10/295,027  
PRIORITY FILING DATE: 2002-11-13  
PRIORITY FILING DATE: 2000-09-15  
PRIORITY FILING DATE: 2000-09-15  
PRIORITY FILING DATE: 2001-11-13  
PRIORITY FILING DATE: 2001-11-13  
PRIORITY FILING DATE: 2001-11-15  
PRIORITY FILING DATE: 2001-11-15  
PRIORITY FILING DATE: 2001-11-21  
PRIORITY FILING DATE: 2001-11-21  
PRIORITY FILING DATE: 2001-11-29

PRIOR APPLICATION NUMBER: US 60/340,376  
PRIOR FILING DATE: 2001-12-14  
PRIOR APPLICATION NUMBER: US 60/347,211  
PRIOR FILING DATE: 2002-01-08  
PRIOR APPLICATION NUMBER: US 60/347,349  
PRIOR FILING DATE: 2002-01-10  
PRIOR APPLICATION NUMBER: US 60/355,250  
PRIOR FILING DATE: 2002-02-08  
PRIOR APPLICATION NUMBER: US 60/356,714  
PRIOR FILING DATE: 2002-02-13  
REMAINING PRIOR APPLICATION DATA REMOVED - See File Wrapper or PALM.  
NUMBER OF SEQ ID NOS: 1386  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 574  
LENGTH: 454  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-10-295-027-574

Query Match 100.0%; Score 2351; DB 15; Length 454;  
Best Local Similarity 100.0%; Pred. No. 8e-222;  
Matches 454; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 1 MESISMWSPKSLSETCLPENGINGIDARKVTGVIGSGDPFASLTIRLRGCVHYVIGS 60
DB 1 MESISMWSPKSLSETCLPENGINGIDARKVTGVIGSGDPFASLTIRLRGCVHYVIGS 60
QY 61 RNPKFASFPFHVDYTHEDALTKTNIIFVAIHREHYTSLMDRLRLVKGKILIDVSNM 120
DB 61 RNPKFASFPFHVDYTHEDALTKTNIIFVAIHREHYTSLMDRLRLVKGKILIDVSNM 120
QY 121 RINQVPSNAEYIASLFPDLSLYKGFNVSAWALQGPDAQRQVYICSNNOARQOYIE 180
DB 121 RINQVPSNAEYIASLFPDLSLYKGFNVSAWALQGPDAQRQVYICSNNOARQOYIE 180
QY 181 LARQNFIPIDGSLSSAREIENLPLRLFTLMRGVVAVVAISLATEFFLYSFVADVHPYA 240
DB 181 LARQNFIPIDGSLSSAREIENLPLRLFTLMRGVVAVVAISLATEFFLYSFVADVHPYA 240
QY 241 RNOQSDFYKIPFIEVNTKLPYVAITLLSLVYLAGLAAAYQLYGKRRPFWLSTWLO 300
DB 241 RNOQSDFYKIPFIEVNTKLPYVAITLLSLVYLAGLAAAYQLYGKRRPFWLSTWLO 300
QY 301 CRKQGLISFPFAMVAVASLCLPMRSEERYFLNMAVQOVANIENTSNWEEVWRIEMY 360
DB 301 CRKQGLISFPFAMVAVASLCLPMRSEERYFLNMAVQOVANIENTSNWEEVWRIEMY 360
QY 361 ISFGIMSLGSLAVTSIPSVSNALNMBREPSFIQSTLGVALISTFHVLIYGMKRAFE 420
DB 361 ISFGIMSLGSLAVTSIPSVSNALNMBREPSFIQSTLGVALISTFHVLIYGMKRAFE 420
QY 421 EBYRFFYTPNPFVLAIVLPSIYVLDLQCRYPD 454
DB 421 EBYRFFYTPNPFVLAIVLPSIYVLDLQCRYPD 454
```

RESULT 5  
US-10-392-190-2

Sequence 2, Application US/10392190  
Publication No. US2004000598A1

GENERAL INFORMATION:

APPLICANT: Devaux, Brigitte  
APPLICANT: Eberhard, David  
APPLICANT: Goddard, Audrey  
APPLICANT: Godowski, Paul J.  
APPLICANT: Grimaldi, J. Christopher  
APPLICANT: Hillan, Kenneth J.  
APPLICANT: Watanabe, Colin K.  
APPLICANT: Wood, William I.  
APPLICANT: Yasura, Daniel I.  
APPLICANT: Zhang, Zemin  
TITLE OF INVENTION: PUMPCN COMPOSITIONS AND USES THEREOF  
FILE REFERENCE: P2994R1C1-US

CURRENT APPLICATION NUMBER: US/10/392,190  
 CURRENT FILING DATE: 2003-03-19  
 PRIOR APPLICATION NUMBER: US 60/235,451  
 PRIOR FILING DATE: 2000-09-26  
 PRIOR APPLICATION NUMBER: PCT/US01/30290  
 PRIOR FILING DATE: 2001-09-26  
 NUMBER OF SEQ ID NOS: 14  
 SEQ ID NO 2  
 LENGTH: 454  
 TYPE: PRT  
 ORGANISM: Homo Sapien  
 US-10-392-190-2

Query Match 100.0%; Score 2351; DB 15; Length 454;  
 Best Local Similarity 100.0%; Pred. No. 222; Indels 0; Gaps 0;  
 Matches 454; Conservative 0; Mismatches 0;

QY 1 MESISMWSPKSLSETCLPNGINGIDARKVTVGVIGSGDPFASKLTIRLRGCHVTVIGS 60  
 DB 1 MESISMWSPKSLSETCLPNGINGIDARKVTVGVIGSGDPFASKLTIRLRGCHVTVIGS 60  
 QY 61 RNKRFASEFPFHVDVTHHEDALTKTNIIFVAIHREHYTSLMDLRHLVVKILIDVSNM 120  
 DB 61 RNKRFASEFPFHVDVTHHEDALTKTNIIFVAIHREHYTSLMDLRHLVVKILIDVSNM 120  
 QY 121 RINQYPSNAEYLAFLPDSLIYKGFNVVSAMALQGPDKASROYICSNNIQARQOYIE 180  
 DB 121 RINQYPSNAEYLAFLPDSLIYKGFNVVSAMALQGPDKASROYICSNNIQARQOYIE 180  
 QY 181 LARQNLPIPDIGLSASAREIENLPRLFTLMGPVVAVSLATFFFLYSFVRDVIHPYA 240  
 DB 181 LARQNLPIPDIGLSASAREIENLPRLFTLMGPVVAVSLATFFFLYSFVRDVIHPYA 240  
 QY 241 RNQOSDPYKPIPIELVNTLPIVAILTLSTVYLAGLAAAYQLYGTYRRFPFMLETMO 300  
 DB 241 RNQOSDPYKPIPIELVNTLPIVAILTLSTVYLAGLAAAYQLYGTYRRFPFMLETMO 300  
 QY 301 CRKQLGLSFFPMVAVAYSLCLPMRSERYLFLNMAVQOVANINSENEEVRLEMY 360  
 DB 301 CRKQLGLSFFPMVAVAYSLCLPMRSERYLFLNMAVQOVANINSENEEVRLEMY 360  
 QY 361 ISFGIMSLGLSLAATVTSISVSANALNMRFEFSIOSTLGVALLISTFVLLYGMKRAFE 420  
 DB 361 ISFGIMSLGLSLAATVTSISVSANALNMRFEFSIOSTLGVALLISTFVLLYGMKRAFE 420  
 QY 421 EERYRFTYPPNFVALVLPISVILDLQLCRYPD 454  
 DB 421 EERYRFTYPPNFVALVLPISVILDLQLCRYPD 454

RESULT 6  
 US-10-104-047-3155  
 Sequence 3155, Application US/10104047  
 Publication No. US20030236392A1  
 GENERAL INFORMATION:  
 APPLICANT: HELIX RESEARCH INSTITUTE  
 TITLE OF INVENTION: No. US20030236392A1 full length cDNA  
 FILE REFERENCE: H1-A0105  
 CURRENT APPLICATION NUMBER: US/10/104,047  
 CURRENT FILING DATE: 2002-09-25  
 PRIOR APPLICATION NUMBER:  
 PRIOR FILING DATE:  
 NUMBER OF SEQ ID NOS: 4096  
 SOFTWARE: Patent Ver. 2.1  
 SEQ ID NO 3155  
 LENGTH: 454  
 TYPE: PRT  
 ORGANISM: Homo sapiens  
 US-10-104-047-3155

Query Match 99.9%; Score 2348; DB 15; Length 454;  
 Best Local Similarity 99.8%; Pred. No. 1,66-221; Indels 0; Gaps 0;  
 Matches 453; Conservative 1; Mismatches 0;

QY 1 MESISMWSPKSLSETCLPNGINGIDARKVTVGVIGSGDPFASKLTIRLRGCHVTVIGS 60  
 DB 1 MESISMWSPKSLSETCLPNGINGIDARKVTVGVIGSGDPFASKLTIRLRGCHVTVIGS 60  
 QY 61 RNKRFASEFPFHVDVTHHEDALTKTNIIFVAIHREHYTSLMDLRHLVVKILIDVSNM 120  
 DB 61 RNKRFASEFPFHVDVTHHEDALTKTNIIFVAIHREHYTSLMDLRHLVVKILIDVSNM 120  
 QY 121 RINQYPSNAEYLAFLPDSLIYKGFNVVSAMALQGPDKASROYICSNNIQARQOYIE 180  
 DB 121 RINQYPSNAEYLAFLPDSLIYKGFNVVSAMALQGPDKASROYICSNNIQARQOYIE 180  
 QY 181 LARQNLPIPDIGLSASAREIENLPRLFTLMGPVVAVSLATFFFLYSFVRDVIHPYA 240  
 DB 181 LARQNLPIPDIGLSASAREIENLPRLFTLMGPVVAVSLATFFFLYSFVRDVIHPYA 240  
 QY 241 RNQOSDPYKPIPIELVNTLPIVAILTLSTVYLAGLAAAYQLYGTYRRFPFMLETMO 300  
 DB 241 RNQOSDPYKPIPIELVNTLPIVAILTLSTVYLAGLAAAYQLYGTYRRFPFMLETMO 300  
 QY 301 CRKQLGLSFFPMVAVAYSLCLPMRSERYLFLNMAVQOVANINSENEEVRLEMY 360  
 DB 301 CRKQLGLSFFPMVAVAYSLCLPMRSERYLFLNMAVQOVANINSENEEVRLEMY 360  
 QY 361 ISFGIMSLGLSLAATVTSISVSANALNMRFEFSIOSTLGVALLISTFVLLYGMKRAFE 420  
 DB 361 ISFGIMSLGLSLAATVTSISVSANALNMRFEFSIOSTLGVALLISTFVLLYGMKRAFE 420  
 QY 421 EERYRFTYPPNFVALVLPISVILDLQLCRYPD 454  
 DB 421 EERYRFTYPPNFVALVLPISVILDLQLCRYPD 454

RESULT 7  
 US-09-802-520-1  
 Sequence 1, Application US/09802520  
 Publication No. US20020187472A1  
 GENERAL INFORMATION:  
 APPLICANT: Lal, Preeti  
 APPLICANT: Farris, Mary  
 APPLICANT: Chen, Hui-Mei  
 APPLICANT: Ison, Craig H.  
 TITLE OF INVENTION: STEAP-RELATED PROTEIN  
 FILE REFERENCE: PC-0037 US  
 CURRENT APPLICATION NUMBER: US/09/802,520  
 CURRENT FILING DATE: 2001-03-09  
 NUMBER OF SEQ ID NOS: 11  
 SOFTWARE: PERL Program  
 SEQ ID NO 1  
 LENGTH: 490  
 TYPE: PRT  
 ORGANISM: Homo sapiens  
 FEATURE:  
 NAME/KEY: misc feature  
 OTHER INFORMATION: Incyte ID No. US20020187472A1 7492448CD1  
 US-09-802-520-1

Query Match 97.6%; Score 2294; DB 9; Length 490;  
 Best Local Similarity 100.0%; Pred. No. 3,56-216; Indels 0; Gaps 0;  
 Matches 444; Conservative 0; Mismatches 0;

QY 1 MESISMWSPKSLSETCLPNGINGIDARKVTVGVIGSGDPFASKLTIRLRGCHVTVIGS 60  
 DB 1 MESISMWSPKSLSETCLPNGINGIDARKVTVGVIGSGDPFASKLTIRLRGCHVTVIGS 60  
 QY 61 RNKRFASEFPFHVDVTHHEDALTKTNIIFVAIHREHYTSLMDLRHLVVKILIDVSNM 120  
 DB 61 RNKRFASEFPFHVDVTHHEDALTKTNIIFVAIHREHYTSLMDLRHLVVKILIDVSNM 120  
 QY 121 RINQYPSNAEYLAFLPDSLIYKGFNVVSAMALQGPDKASROYICSNNIQARQOYIE 180  
 DB 121 RINQYPSNAEYLAFLPDSLIYKGFNVVSAMALQGPDKASROYICSNNIQARQOYIE 180

```
QY 181 LARQNFIPIDIGSSSAREIENLPLRLFTLRGVPVVAISLATFFLYSFVRDIHPYA 240
Db 181 LARQNFIPIDIGSSSAREIENLPLRLFTLRGVPVVAISLATFFLYSFVRDIHPYA 240
QY 241 RNQOSDFYKIPLEIYNKTLPIVAITLLSLVYLAGLLAAAYQLYGTXRRPPEMLWTWQ 300
Db 241 RNQOSDFYKIPLEIYNKTLPIVAITLLSLVYLAGLLAAAYQLYGTXRRPPEMLWTWQ 300
QY 301 CRKQGLISFFPAMVAVASLCLPMRBSERYFLNMAVQOVANINENSMBEVRIMY 360
Db 301 CRKQGLISFFPAMVAVASLCLPMRBSERYFLNMAVQOVANINENSMBEVRIMY 360
QY 361 ISFGIMSLGILLSLAVTSIPSVSNALNMBEFSFIQSTLGYALLSTPHVLIYGMKRAFE 420
Db 361 ISFGIMSLGILLSLAVTSIPSVSNALNMBEFSFIQSTLGYALLSTPHVLIYGMKRAFE 420
QY 421 EBYRFTYTPPNFVALVLPISIVIL 444
Db 421 EBYRFTYTPPNFVALVLPISIVIL 444
```

## RESULT 8

```
US-10-239-607-14
; Sequence 14, Application US/10239607
; Publication No. US20030219761A1
; GENERAL INFORMATION:
; APPLICANT: Saatchioglul, Fahri
; TITLE OF INVENTION: No. US20030219761A1el Prostate-Specific or
; TITLE OF INVENTION: Testis-Specific Nucleic Acid Molecules, Polypeptides, and
; FILE REFERENCE: 50218/003002
; CURRENT APPLICATION NUMBER: US/10/239, 607
; PRIOR FILING DATE: 2002-09-24
; PRIOR APPLICATION NUMBER: PCT/US01/09410
; PRIOR FILING DATE: 2001-03-23
; PRIOR APPLICATION NUMBER: US 60/191,929
; NUMBER OF SEQ ID NOS: 73
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14
; LENGTH: 490
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-239-607-14
```

Query Match Best Local Similarity 97.4%; Score 2290; DB 15; Length 490;

Matches 443; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
QY 1 MESISMGSPKSLSETCLPNGINGIKDARKVTGVIGSGDPFASKLITRLIRCGYHVYIGS 60
Db 1 MESISMGSPKSLSETCLPNGINGIKDARKVTGVIGSGDPFASKLITRLIRCGYHVYIGS 60
QY 61 RNPKFASSEFPFHVVDVTHEEDALTNTNIIFVAIHREHTSLMDRLHLVGKILLIDVSNM 120
Db 61 RNPKFASSEFPFHVVDVTHEEDALTNTNIIFVAIHREHTSLMDRLHLVGKILLIDVSNM 120
QY 121 RINQYPSNAEYIASLFPDLSLVKGFNVVSAMALQGPDKASROYICSNNTQARQOYIE 180
Db 121 RINQYPSNAEYIASLFPDLSLVKGFNVVSAMALQGPDKASROYICSNNTQARQOYIE 180
QY 181 LARQNFIPIDIGSSSAREIENLPLRLFTLRGVPVVAISLATFFLYSFVRDIHPYA 240
Db 181 LARQNFIPIDIGSSSAREIENLPLRLFTLRGVPVVAISLATFFLYSFVRDIHPYA 240
QY 241 RNQOSDFYKIPLEIYNKTLPIVAITLLSLVYLAGLLAAAYQLYGTXRRPPEMLWTWQ 300
Db 241 RNQOSDFYKIPLEIYNKTLPIVAITLLSLVYLAGLLAAAYQLYGTXRRPPEMLWTWQ 300
QY 301 CRKQGLISFFPAMVAVASLCLPMRBSERYFLNMAVQOVANINENSMBEVRIMY 360
Db 301 CRKQGLISFFPAMVAVASLCLPMRBSERYFLNMAVQOVANINENSMBEVRIMY 360
```

```
QY 361 ISFGIMSLGILLSLAVTSIPSVSNALNMBEFSFIQSTLGYALLSTPHVLIYGMKRAFE 420
Db 361 ISFGIMSLGILLSLAVTSIPSVSNALNMBEFSFIQSTLGYALLSTPHVLIYGMKRAFE 420
QY 421 EBYRFTYTPPNFVALVLPISIVIL 444
Db 421 EBYRFTYTPPNFVALVLPISIVIL 444
```

## RESULT 9

```
US-10-239-607-32
; Sequence 32, Application US/10239607
; Publication No. US20030219761A1
; GENERAL INFORMATION:
; APPLICANT: Saatchioglul, Fahri
; TITLE OF INVENTION: No. US20030219761A1el Prostate-Specific or
; TITLE OF INVENTION: Testis-Specific Nucleic Acid Molecules, Polypeptides, and
; FILE REFERENCE: 50218/003002
; CURRENT APPLICATION NUMBER: US/10/239, 607
; PRIOR FILING DATE: 2002-09-24
; PRIOR APPLICATION NUMBER: PCT/US01/09410
; PRIOR FILING DATE: 2001-03-23
; PRIOR APPLICATION NUMBER: US 60/191,929
; NUMBER OF SEQ ID NOS: 73
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 32
; LENGTH: 419
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-239-607-32
```

Query Match Best Local Similarity 86.6%; Score 2036; DB 15; Length 419;

Matches 394; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
QY 1 MESISMGSPKSLSETCLPNGINGIKDARKVTGVIGSGDPFASKLITRLIRCGYHVYIGS 60
Db 1 MESISMGSPKSLSETCLPNGINGIKDARKVTGVIGSGDPFASKLITRLIRCGYHVYIGS 60
QY 61 RNPKFASSEFPFHVVDVTHEEDALTNTNIIFVAIHREHTSLMDRLHLVGKILLIDVSNM 120
Db 61 RNPKFASSEFPFHVVDVTHEEDALTNTNIIFVAIHREHTSLMDRLHLVGKILLIDVSNM 120
QY 121 RINQYPSNAEYIASLFPDLSLVKGFNVVSAMALQGPDKASROYICSNNTQARQOYIE 180
Db 121 RINQYPSNAEYIASLFPDLSLVKGFNVVSAMALQGPDKASROYICSNNTQARQOYIE 180
QY 181 LARQNFIPIDIGSSSAREIENLPLRLFTLRGVPVVAISLATFFLYSFVRDIHPYA 240
Db 181 LARQNFIPIDIGSSSAREIENLPLRLFTLRGVPVVAISLATFFLYSFVRDIHPYA 240
QY 241 RNQOSDFYKIPLEIYNKTLPIVAITLLSLVYLAGLLAAAYQLYGTXRRPPEMLWTWQ 300
Db 241 RNQOSDFYKIPLEIYNKTLPIVAITLLSLVYLAGLLAAAYQLYGTXRRPPEMLWTWQ 300
QY 301 CRKQGLISFFPAMVAVASLCLPMRBSERYFLNMAVQOVANINENSMBEVRIMY 360
Db 301 CRKQGLISFFPAMVAVASLCLPMRBSERYFLNMAVQOVANINENSMBEVRIMY 360
QY 361 ISFGIMSLGILLSLAVTSIPSVSNALNMBEFSFIQ 395
Db 361 ISFGIMSLGILLSLAVTSIPSVSNALNMBEFSFIQ 395
```

## RESULT 10

```
US-10-182-951-6
; Sequence 6, Application US/10182951
; Publication No. US20030138895A1
; GENERAL INFORMATION:
; APPLICANT: INCYTE GENOMICS, INC.
```



```
APPLICANT: Saaticoglu, Fahri
; TITLE OF INVENTION: No. US20030219761A1el Prostate-Specific or
; TITLE OF INVENTION: Testis-Specific Nucleic Acid Molecules, Polypeptides, and
; TITLE OF INVENTION: Diagnostic and Therapeutic Methods
; FILE REFERENCE: 50218/003002
; CURRENT APPLICATION NUMBER: US/10/239,607
; CURRENT FILING DATE: 2002-09-24
; PRIOR APPLICATION NUMBER: PCT/US01/09410
; PRIOR FILING DATE: 2001-03-23
; PRIOR APPLICATION NUMBER: US 60/191,929
; PRIOR FILING DATE: 2000-03-24
; NUMBER OF SEQ ID NOS: 73
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 36
; LENGTH: 488
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-239-607-36
```

```
Query Match      54.0%; Score 1269; DB 15; Length 488;
Best Local Similarity 54.3%; Pred. No. 1,3e-115;
Matches 241; Conservative 83; Mismatches 108; Indels 12; Gaps 4;
```

```
QY 3 SISMGSPKSLSETCLPENGINGIKDARKVTYGVIGSGDFAKSLTIRLCGYHVIGSN 62
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 11 SHLVDSDSLAK--VPD-----EAPK--VGILSGDFARSLATRLVSGGFKVVGSRN 60
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 63 PKFAEFPPHYVDVTHEDALTKNIIFVAIHREHTSLMDRLHVGKILIDVSN--NM 120
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 61 PKTKRLYPSAAQVTFQBEAVSPVITFAVREHYSLSLSLQAGKILIDVSNPTQ 120
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 121 RINQVPSNAEYLASLPDLSLVKGFNVVSAMALQGPDKASROYVIGSNNOIQAQVIE 180
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 121 EHLQRESNAEYLASLPDLSLVKGFNVVSAMALQGPDKASROYVIGSNNOIQAQVIE 180
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 181 LARQNLFPIDGLSLSARIEINLPRLFTLMRGPVVAISLATEFFLYSPVDVHPYA 240
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 181 MALMGFPMVDMGSLASAMEVAMPRLPAMKVPFTLLALGLFVCFYAVNFRDVLQPYV 240
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 241 RNOQSDFFKIPLEIVNKTLPVAILTSLVYLAGLAAVQLYGTKRFRFPMLTQ 300
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 241 QESQNKFFKLPSVAVNTLPCVAVYLSVYLPGLVLAALQRRGTYKQRPDMIDHMLQ 300
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 301 CRKQGLISFFPAMVAVSYLCPMRSERYLFNMAVQOVHANIENSNWEEVARIENY 360
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 301 HRKQGLISFFPAMVAVSYLCPMRSERYLFNMAVQOVHANIENSNWEEVARIENY 360
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 361 ISFGIMSLGLSLAVTSIPSVSNALNRESEFIOSTLGYVALLISTFHVLYGMKRAFE 420
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 361 LSLGLVLTSLAVTSIPSVSNALNRESEFIOSTLGYVALLISTFHVLYGMKRAFE 420
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 421 EERYRFTPPNFVTLVPSIVIL 444
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 421 ESRYKFLPPTFTLLVPCVIL 444
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
```

```
RESULT 13
US-10-239-607-38
```

```
; Sequence 38, Application US/10239607
```

```
; Publication No. US20030219761A1
```

```
; GENERAL INFORMATION:
```

```
; APPLICANT: Saaticoglu, Fahri
```

```
; TITLE OF INVENTION: No. US20030219761A1el Prostate-Specific or
```

```
; TITLE OF INVENTION: Testis-Specific Nucleic Acid Molecules, Polypeptides, and
```

```
; TITLE OF INVENTION: Diagnostic and Therapeutic Methods
```

```
; FILE REFERENCE: 50218/003002
```

```
; CURRENT APPLICATION NUMBER: US/10/239,607
```

```
; CURRENT FILING DATE: 2002-09-24
```

```
; PRIOR APPLICATION NUMBER: PCT/US01/09410
```

```
; PRIOR FILING DATE: 2001-03-23
```

```
; PRIOR APPLICATION NUMBER: US 60/191,929
```

```
; PRIOR FILING DATE: 2000-03-24
```

```
; NUMBER OF SEQ ID NOS: 73
```

```
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 38
; LENGTH: 488
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-239-607-38
```

```
Query Match      54.0%; Score 1269; DB 15; Length 488;
Best Local Similarity 54.3%; Pred. No. 1,3e-115;
Matches 241; Conservative 83; Mismatches 108; Indels 12; Gaps 4;
```

```
QY 3 SISMGSPKSLSETCLPENGINGIKDARKVTYGVIGSGDFAKSLTIRLCGYHVIGSN 62
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 11 SHLVDSDSLAK--VPD-----EAPK--VGILSGDFARSLATRLVSGGFKVVGSRN 60
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 63 PKFAEFPPHYVDVTHEDALTKNIIFVAIHREHTSLMDRLHVGKILIDVSN--NM 120
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 61 PKTKRLYPSAAQVTFQBEAVSPVITFAVREHYSLSLSLQAGKILIDVSNPTQ 120
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 121 RINQVPSNAEYLASLPDLSLVKGFNVVSAMALQGPDKASROYVIGSNNOIQAQVIE 180
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 121 EHLQRESNAEYLASLPDLSLVKGFNVVSAMALQGPDKASROYVIGSNNOIQAQVIE 180
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 181 LARQNLFPIDGLSLSARIEINLPRLFTLMRGPVVAISLATEFFLYSPVDVHPYA 240
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 181 MALMGFPMVDMGSLASAMEVAMPRLPAMKVPFTLLALGLFVCFYAVNFRDVLQPYV 240
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 241 RNOQSDFFKIPLEIVNKTLPVAILTSLVYLAGLAAVQLYGTKRFRFPMLTQ 300
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 241 QESQNKFFKLPSVAVNTLPCVAVYLSVYLPGLVLAALQRRGTYKQRPDMIDHMLQ 300
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 301 CRKQGLISFFPAMVAVSYLCPMRSERYLFNMAVQOVHANIENSNWEEVARIENY 360
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 301 HRKQGLISFFPAMVAVSYLCPMRSERYLFNMAVQOVHANIENSNWEEVARIENY 360
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 361 ISFGIMSLGLSLAVTSIPSVSNALNRESEFIOSTLGYVALLISTFHVLYGMKRAFE 420
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 361 LSLGLVLTSLAVTSIPSVSNALNRESEFIOSTLGYVALLISTFHVLYGMKRAFE 420
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 421 EERYRFTPPNFVTLVPSIVIL 444
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 421 ESRYKFLPPTFTLLVPCVIL 444
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
```

```
RESULT 14
US-09-965-529-21
```

```
; Sequence 21, Application US/09965529
```

```
; Publication No. US20020182671A1
```

```
; GENERAL INFORMATION:
```

```
; APPLICANT: Lal, Preeti
```

```
; APPLICANT: YUE, Henry
```

```
; APPLICANT: TANG, Y. Tom
```

```
; APPLICANT: BANDMAN, Olga
```

```
; APPLICANT: BUREFORD, Neil
```

```
; APPLICANT: AZIMZAI, Yalda
```

```
; APPLICANT: BAUGHN, Mariah R.
```

```
; APPLICANT: LU, Dying Aina M.
```

```
; APPLICANT: PATTERSON, Chandra
```

```
; TITLE OF INVENTION: MEMBRANE ASSOCIATED PROTEINS
```

```
; FILE REFERENCE: PF-0731 USA
```

```
; CURRENT APPLICATION NUMBER: US/09/965,529
```

```
; CURRENT FILING DATE: 2001-09-26
```

```
; PRIOR APPLICATION NUMBER: 60/149,641; 60/164,203; PCT/US00/22315
```

```
; PRIOR FILING DATE: 1999-08-17; 1999-11-09; 2000-08-14
```

```
; NUMBER OF SEQ ID NOS: 74
```

```
; SOFTWARE: PERL Program
```

```
; SEQ ID NO 21
```

```
; LENGTH: 459
```

```
; TYPE: PRT
```

```
; ORGANISM: Homo sapiens
```

```
; NAME/KEY: misc.feature
```

```
; OTHER INFORMATION: Incyte ID No. US20020182671A1 1859305CD1
```



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 1, 2004, 19:57:15 ; Search time 83 Seconds

(without alignments)  
1725.847 Million cell updates/sec

Title: US-09-455-486-6

Perfect score: 2351

Sequence: 1 MESISMWSPKSLSETCLPN.....ALVPSIVILDLLQCRYPD 454

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL\_25:\*  
1: sp\_archaea:\*  
2: sp\_bacteria:\*  
3: sp\_fungi:\*  
4: sp\_human:\*  
5: sp\_invertebrate:\*  
6: sp\_mammal:\*  
7: sp\_rhco:\*  
8: sp\_organelle:\*  
9: sp\_phage:\*  
10: sp\_plant:\*  
11: sp\_rudent:\*  
12: sp\_virus:\*  
13: sp\_vertebrate:\*  
14: sp\_unclassified:\*  
15: sp\_virus:\*  
16: sp\_bacteriaph:\*  
17: sp\_archaeap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2290	97.4	490	4 Q81UE7	Q81UE7 homo sapien
2	2284	97.2	490	4 Q8NFT2	Q8NFT2 homo sapien
3	2239.5	95.3	489	11 Q8BMB6	Q8BMB6 mus musculu
4	1318	56.1	488	11 Q8CSP0	Q8CSP0 mus musculu
5	1316	56.0	488	11 Q99P41	Q99P41 ratcus norv
6	1316	56.0	488	11 Q8CIS9	Q8CIS9 mus musculu
7	1316	56.0	526	11 Q80ZF3	Q80ZF3 mus musculu
8	1306	55.6	514	11 Q924Z1	Q924Z1 mus musculu
9	1272	54.1	488	4 Q8SFF6	Q8SFF6 homo sapien
10	1272	54.1	498	4 Q7Z389	Q7Z389 homo sapien
11	1269	53.3	488	4 Q9NV55	Q9NV55 homo sapien
12	1252.5	53.0	487	4 Q8NEM6	Q8NEM6 homo sapien
13	1245.5	53.0	487	4 Q8TF03	Q8TF03 homo sapien
14	1074.5	45.7	456	4 Q8TDP3	Q8TDP3 homo sapien
15	1064.5	45.3	470	11 Q923B6	Q923B6 mus musculu
16	1057.5	45.0	470	11 Q91W31	Q91W31 mus musculu

17	1041.5	44.3	474	11 Q91ZB8	Q91ZB8 mus musculu
18	956.5	40.7	464	11 Q7TP88	Q7TP88 ratcus norv
19	728	31.0	338	6 Q9GL50	Q9GL50 sus scrofa
20	715	30.4	339	11 Q924Z2	Q924Z2 mus musculu
21	707	30.1	339	11 Q9CWR7	Q9CWR7 mus musculu
22	704	29.9	339	11 Q924J9	Q924J9 mus musculu
23	628	26.7	283	4 Q9H5R1	Q9H5R1 homo sapien
24	541	23.0	283	4 Q8WMB0	Q8WMB0 homo sapien
25	324	13.8	143	4 Q8H7Y1	Q8H7Y1 homo sapien
26	171.5	7.3	208	16 Q8NTU6	Q8NTU6 corynebacte
27	170	7.2	208	16 Q8XOS2	Q8XOS2 ralsstonia s
28	164.5	7.0	234	16 Q82AX0	Q82AX0 streptomyc
29	160	6.8	198	16 Q930K7	Q930K7 rhizobium m
30	159.5	6.8	211	16 Q8TK44	Q8TK44 anabaena sp
31	156.5	6.7	229	16 Q9TKR6	Q9TKR6 streptomyc
32	154.5	6.6	235	2 Q8GFG4	Q8GFG4 rhodococcus
33	151	6.4	200	16 Q98C62	Q98C62 rhizobium 1
34	149.5	6.4	218	2 Q8GFG3	Q8GFG3 rhodococcus
35	149.5	6.4	221	16 Q82Q14	Q82Q14 streptomyc
36	147.5	6.3	226	2 Q8GK65	Q8GK65 rhodococcus
37	147	6.3	213	17 Q930S9	Q930S9 archaeoglob
38	144.5	6.1	226	2 Q9AH05	Q9AH05 rhodococcus
39	143.5	6.1	223	17 Q8TYQ9	Q8TYQ9 methanopyru
40	142	6.0	212	17 Q29370	Q29370 archaeoglob
41	142	6.0	224	1 O59661	O59661 methanobact
42	137.5	5.8	199	2 Q9F418	Q9F418 mycobacteri
43	136.5	5.8	217	2 Q83VC1	Q83VC1 streptomyc
44	136.5	5.8	217	16 Q9UN78	Q9UN78 streptomyc
45	135.5	5.8	203	2 Q8GFG1	Q8GFG1 rhodococcus

## ALIGNMENTS

RESULT 1	CE1UE7	PRELIMINARY;	PRT;	490 AA.
ID	Q81UE7			
AC	Q81UE7			
DT	01-MAR-2003 (T-EMBLrel. 23, Created)			
DT	01-MAR-2003 (T-EMBLrel. 23, Last sequence update)			
DT	01-UTN-2003 (T-EMBLrel. 24, Last annotation update)			
DE	STAMPI.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Homindae; Homo.			
OX	NCBI_Taxid=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Prostate;			
RX	MEDLINE=22229309; PubMed=12095985;			
RA	Korkmaz K.S., Elbl C.C., Korkmaz C.G., Ioda M., Hager G.L.,			
RA	Saetcioglu F.;			
RT	"Molecular cloning and characterization of STAMPI, a highly prostate			
RT	specific six-trans-membrane protein that is overexpressed in prostate			
RT	cancer."			
RL	J. Biol. Chem. 277:36689-36696(2002).			
DR	EMBL; AY008445; AAG32149.1; -			
DR	GO; GO:0006118; P:electon transport; IBA.			
DR	InterPro; IPR004455; NADPoxred.F420.			
DR	Pfam; PF03807; F420_oxidored; I			
SQ	SEQUENCE 490 AA; 56027 MW; D1971A84880F7E51 CRC64;			
Query Match	97.4%; Score 2290; DB 4; Length 490;			
Best Local Similarity	99.8%; Pred. No. 4.1e-176;			
Matches 443; Conservative	0; Mismatches 1; Indels 0; Gaps 0;			
QY	1 MESISMWSPKSLSETCLPNNGINGIKDARKVGVIGSDPFAKSLTIRIRGCVHVIGS 60			
DB	1 MESISMWSPKSLSETCLPNNGINGIKDARKVGVIGSDPFAKSLTIRIRGCVHVIGS 60			
QY	61 RNPFASEFPFHVVDVTHEDALTKNNIIFVAIHREHYTSLDHLHLVGLKILDIVSNM 120			
DB	61 RNPFASEFPFHVVDVTHEDALTKNNIIFVAIHREHYTSLDHLHLVGLKILDIVSNM 120			



```

QY 121 RINQYPSNAEYLAELPDSLTVKGFNVSAWALQIGPKDASROYVICSNNIOARQOYIE 180
DB 121 RINQYPSNAEYLAELPDSLTVKGFNVSAWALQIGPKDASROYVICSNNIOARQOYIE 180
QY 181 LARQNLNFIPIIDGLSSASAREIENLPLRLFTLMRGVVAVIAISLATFFFLYSFVADVHPYA 240
DB 181 LARQNLNFIPIIDGLSSASAREIENLPLRLFTLMRGVVAVIAISLATFFFLYSFVADVHPYA 240
QY 241 RNOQSDFFYKIPRIEIVNKTLPVIAITLLSLVYLAGLAAAYOLYGTXYRRPFWLETWLQ 300
DB 241 RNOQSDFFYKIPRIEIVNKTLPVIAITLLSLVYLAGLAAAYOLYGTXYRRPFWLETWLQ 300
QY 301 CRKQGLISFFPAMVAVASLCLPMRSEERYFLMAYOOVHANINENSWNEEVRMIEM 360
DB 301 CRKQGLISFFPAMVAVASLCLPMRSEERYFLMAYOOVHANINENSWNEEVRMIEM 360
QY 361 ISFGINSGLSLSLATVTSIPSVSNALNWRSEFSIOSTLGYVALLISTFVLLIYGMKRAFE 420
DB 361 ISFGINSGLSLSLATVTSIPSVSNALNWRSEFSIOSTLGYVALLISTFVLLIYGMKRAFE 420
QY 421 EBYRFPYPPNFVLAIVLPSIVIL 444
DB 421 EBYRFPYPPNFVLAIVLPSIVIL 444
QY 421 EBYRFPYPPNFVLAIVLPSIVIL 444
DB 421 EBYRFPYPPNFVLAIVLPSIVIL 444
RESULT 2
Q8NFT2 PRELIMINARY; PRT; 490 AA.
ID Q8NFT2;
AC Q8NFT2;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Six-transmembrane epithelial antigen of prostate 2.
GN STEAP2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OC NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Prostate;
RA Porfika K.P., Helenius M.A., Visakorpi T.;
RT "Cloning and characterization of a novel six-transmembrane protein
RT STEAP2, down-regulated in androgen-independent prostate cancer.";
RL Submitted (DEC-2001) to the EMBL/Genbank/DBJ databases.
DR EMBL; AF455139; FAN04080.1; -.
DR Gene; HGNC:17885; STEAP2.
DR GO; GO:0016021; C:Integral to membrane; IEA.
DR GO; GO:0006118; P:Electron transport; IEA.
DR InterPro; IPR004455; NADPOxd_F420.
DR Pfam; PF03807; F420_oxidored; I.
KM Transmembrane.
SQ SEQUENCE 490 AA; 55961 MW; F2E8C30CDACEF81 CRC64;
Query Match 97.2%; Score 2284; DB 4; Length 490;
Best Local Similarity 99.8%; Pred. No. 1.2e-175;
Matches 443; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 MESISMWSPKSLSETCLPENGINGIDARVTVGVIGSGDFAKSLTIRLRGCVHYVIGS 60
DB 1 MESISMWSPKSLSETCLPENGINGIDARVTVGVIGSGDFAKSLTIRLRGCVHYVIGS 60
QY 61 RNPKEASEFPFPHVADVTHEDALTKNNIIFVAIHREHYTSLMDLRHLVKGKILIDVSNM 120
DB 61 RNPKEASEFPFPHVADVTHEDALTKNNIIFVAIHREHYTSLMDLRHLVKGKILIDVSNM 120
QY 121 RINQYPSNAEYLAELPDSLTVKGFNVSAWALQIGPKDASROYVICSNNIOARQOYIE 180
DB 121 RINQYPSNAEYLAELPDSLTVKGFNVSAWALQIGPKDASROYVICSNNIOARQOYIE 180
QY 181 LARQNLNFIPIIDGLSSASAREIENLPLRLFTLMRGVVAVIAISLATFFFLYSFVADVHPYA 240
DB 181 LARQNLNFIPIIDGLSSASAREIENLPLRLFTLMRGVVAVIAISLATFFFLYSFVADVHPYA 240

```

```

DB 181 LARQNLNFIPIIDGLSSASAREIENLPLRLFTLMRGVVAVIAISLATFFFLYSFVADVHPYA 240
QY 241 RNOQSDFFYKIPRIEIVNKTLPVIAITLLSLVYLAGLAAAYOLYGTXYRRPFWLETWLQ 300
DB 241 RNOQSDFFYKIPRIEIVNKTLPVIAITLLSLVYLAGLAAAYOLYGTXYRRPFWLETWLQ 300
QY 301 CRKQGLISFFPAMVAVASLCLPMRSEERYFLMAYOOVHANINENSWNEEVRMIEM 360
DB 301 CRKQGLISFFPAMVAVASLCLPMRSEERYFLMAYOOVHANINENSWNEEVRMIEM 360
QY 361 ISFGINSGLSLSLATVTSIPSVSNALNWRSEFSIOSTLGYVALLISTFVLLIYGMKRAFE 420
DB 361 ISFGINSGLSLSLATVTSIPSVSNALNWRSEFSIOSTLGYVALLISTFVLLIYGMKRAFE 420
QY 421 EBYRFPYPPNFVLAIVLPSIVIL 444
DB 421 EBYRFPYPPNFVLAIVLPSIVIL 444
QY 421 EBYRFPYPPNFVLAIVLPSIVIL 444
DB 421 EBYRFPYPPNFVLAIVLPSIVIL 444
RESULT 3
Q8BWB6 PRELIMINARY; PRT; 489 AA.
ID Q8BWB6;
AC Q8BWB6;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Weakly similar to tumor suppressor PHFDE.
GN 421358B17RIK.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OC NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Head;
RX MEDLINE=22354683; PubMed=12466851;
RA The FANTOM Consortium;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs.";
RL Nature 420:563-573(2002).
DR EMBL; AK052981; BAC35230.1; -.
DR MGD; MGI:1921301; 492153B817RIK.
DR GO; GO:0006118; P:Electron transport; IEA.
DR InterPro; IPR004455; NADPOxd_F420.
DR Pfam; PF03807; F420_oxidored; I.
SQ SEQUENCE 489 AA; 55760 MW; 98CD63D59DDDF24C CRC64;
Query Match 95.3%; Score 2239.5; DB 11; Length 489;
Best Local Similarity 97.3%; Pred. No. 4.8e-172;
Matches 432; Conservative 9; Mismatches 2; Indels 1; Gaps 1;
QY 1 MESISMWSPKSLSETCLPENGINGIDARVTVGVIGSGDFAKSLTIRLRGCVHYVIGS 60
DB 1 MESISMWSPKSLSETCLPENGINGIDARVTVGVIGSGDFAKSLTIRLRGCVHYVIGS 60
QY 61 RNPKEASEFPFPHVADVTHEDALTKNNIIFVAIHREHYTSLMDLRHLVKGKILIDVSNM 120
DB 61 RNPKEASEFPFPHVADVTHEDALTKNNIIFVAIHREHYTSLMDLRHLVKGKILIDVSNM 120
QY 121 RINQYPSNAEYLAELPDSLTVKGFNVSAWALQIGPKDASROYVICSNNIOARQOYIE 180
DB 121 RINQYPSNAEYLAELPDSLTVKGFNVSAWALQIGPKDASROYVICSNNIOARQOYIE 180
QY 181 LARQNLNFIPIIDGLSSASAREIENLPLRLFTLMRGVVAVIAISLATFFFLYSFVADVHPYA 240
DB 181 LARQNLNFIPIIDGLSSASAREIENLPLRLFTLMRGVVAVIAISLATFFFLYSFVADVHPYA 240
QY 241 RNOQSDFFYKIPRIEIVNKTLPVIAITLLSLVYLAGLAAAYOLYGTXYRRPFWLETWLQ 300
DB 241 RNOQSDFFYKIPRIEIVNKTLPVIAITLLSLVYLAGLAAAYOLYGTXYRRPFWLETWLQ 300
QY 301 CRKQGLISFFPAMVAVASLCLPMRSEERYFLMAYOOVHANINENSWNEEVRMIEM 360
DB 301 CRKQGLISFFPAMVAVASLCLPMRSEERYFLMAYOOVHANINENSWNEEVRMIEM 360

```

```

Db 300 CRKQGLSTFFFAVHVAVSLCLPFRSERLFLNMAVQCHANIENANNEEYTRLEMY 359
Qy 361 ISFGIMSLGLSLAAVSIPIBSVSNLNMREPSFIQSTLGYALLISTPHVLIYGMKRAFE 420
Db 360 ISFGIMSLGLSLAAVSIPIBSVSNLNMREPSFIQSTLGYALLISTPHVLIYGMKRAFA 419
Qy 421 EBYRFPYTPNFVALVPSIVIL 444
Db 420 EBYRFPYTPNFVALVPSIVIL 443

```

## RESULT 4

```

Q8CSF0 PRELIMINARY; PRT; 488 AA.
AC Q8CSF0;
DT 01-MAR-2003 (TREMblrel. 23, Created)
DT 01-MAR-2003 (TREMblrel. 23, Last sequence update)
DT 01-JUN-2003 (TREMblrel. 24, Last annotation update)
DE DUDULIN 2 homology.
GN TSAF6 OR 1010001DOLRIK.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
CX NCBI_TaxId=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Testis;
RX MEDLINE=22354683; PubMed=12466851;
RA The FANTOM Consortium;
RT "Analysis of the mouse transcriptome based on functional annotation of
  60,770 full-length cDNAs."
RL Nature 420:563-573 (2002).
DR EMBL; AK078769; BAC37383.1; -.
DR MGI; MGI:1915678; Tsap6.
DR GO; GO:0006118; P:electron transport; IEA.
DR InterPro; IPR004455; NADPoxred_F420.
DR Pfam; PF03807; F420 oxidored; I.
SQ SEQUENCE 488 AA; 54780 MW; FDFCAF42AE503D11 CRC64;

```

```

Query Match 56.1%; Score 1318; DB 11; Length 488;
Best Local Similarity 56.2%; Pred. No. 8.4e-98;
Matches 248; Conservative 81; Mismatches 100; Indels 12; Gaps 4;
Qy 6 MMGSPKSLSETCTLPNGINGIKDARKYTVGVIGSGDFPKSLTIRLIRCGYHVYVIGSNPKF 65
Db 14 LVSDSGSLAE--VP-----KEAPK--VGILGSDGFARSLATRLVSGSFVVGSGSNPKR 63
Qy 66 ASEFPFHVVDYTHHEDALTNTNIIFFVAIHREHYTSLMDLRLHLYGKILDIVSNMNRIN-- 123
Db 64 TAGLFPBLAQVTFQEEAMSSPEVIFVAVFEHYSSLSGLADQLAGKILVDVSNPTEKRL 123
Qy 124 QYPSNAYLASLPDLSLIVGFNVVSAWALQGPXASROYICSNNTQARQOYIELAR 183
Db 124 QYPSNAYLASLPDLSLIVGFNVVSAWALQGPXASROYICSNNTQARQOYIELAR 183
Qy 124 QYPSNAYLASLPDLSLIVGFNVVSAWALQGPXASROYICSNNTQARQOYIELAR 183
Db 124 QYPSNAYLASLPDLSLIVGFNVVSAWALQGPXASROYICSNNTQARQOYIELAR 183
Qy 184 QINFTPIDLSLSAREIENPLRLFTLMRGPVVAISLATFFLYSPVVDVHYPRANQ 243
Db 184 QINFTPIDLSLSAREIENPLRLFTLMRGPVVAISLATFFLYSPVVDVHYPRANQ 243
Qy 184 AMGSTPILDMGSLASAREVEALPLRLFSMVKVPTLLMGLFVCFTYTNFTRDVLQPIRKO 243
Db 244 QSDPYKPIETVNTKLPVAVATLISLVYLAGLAAVQVLYGKXRRPFWLSTWLOCRK 303
Qy 244 ENKFKYKPLSVNTTLPVAVATLISLVYLAGLAAVQVLYGKXRRPFWLSTWLOCRK 303
Db 244 ENKFKYKPLSVNTTLPVAVATLISLVYLAGLAAVQVLYGKXRRPFWLSTWLOCRK 303
Qy 304 QLGILSFFFAVHVAVSLCLPFRSERLFLNMAVQCHANIENANNEEYTRLEMYISF 363
Db 304 QLGILSFFFAVHVAVSLCLPFRSERLFLNMAVQCHANIENANNEEYTRLEMYISF 363
Qy 364 QIGLISFFFAVHVAVSLCLPFRSERLFLNMAVQCHANIENANNEEYTRLEMYISF 423
Db 364 QIGLISFFFAVHVAVSLCLPFRSERLFLNMAVQCHANIENANNEEYTRLEMYISF 423

```

```

Qy 424 YRFPYTPNFVALVPSIVIL 444
Db 424 YRFPYTPNFVALVPSIVIL 444

```

## RESULT 5

```

Q89P41 PRELIMINARY; PRT; 488 AA.
AC Q89P41;
DT 01-JUN-2001 (TREMblrel. 17, Created)
DT 01-JUN-2001 (TREMblrel. 17, Last sequence update)
DT 01-JUN-2003 (TREMblrel. 24, Last annotation update)
DE Tumor suppressor pHyde.
GN Rattus norvegicus (Rat).
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
CX NCBI_TaxId=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Dunning; TISSUE=Prostatic carcinoma;
RX MEDLINE=20424188; PubMed=10969787;
RA Steiner M.S., Zhang X., Wang Y., Lu Y.;
RT "Growth inhibition of prostate cancer by an adenovirus expressing a
  novel tumor suppressor gene, pHyde."
RL Cancer Res. 60:4419-4425 (2000).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Dunning; TISSUE=Prostatic carcinoma;
RX Lu Y., Rinaldy A.R., Steiner M.S.;
RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC Steiner M.S., Rinaldy A.;
RT "A novel prostate cancer associated gene."
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF335281; AAK0361.1; -.
DR EMBL; AF335281; AAK0361.1; -.
DR GO; GO:0006118; P:electron transport; IEA.
DR InterPro; IPR004455; NADPoxred_F420.
DR Pfam; PF03807; F420 oxidored; I.
SQ SEQUENCE 488 AA; 54640 MW; 59FF07121919FDCB CRC64;

```

```

Query Match 56.0%; Score 1316; DB 11; Length 488;
Best Local Similarity 56.2%; Pred. No. 1.2e-97;
Matches 248; Conservative 81; Mismatches 100; Indels 12; Gaps 4;
Qy 6 MMGSPKSLSETCTLPNGINGIKDARKYTVGVIGSGDFPKSLTIRLIRCGYHVYVIGSNPKF 65
Db 14 LVSDSGSLAE--VP-----KEAPK--VGILGSDGFARSLATRLVSGSFVVGSGSNPKR 63
Qy 66 ASEFPFHVVDYTHHEDALTNTNIIFFVAIHREHYTSLMDLRLHLYGKILDIVSNMNRIN-- 123
Db 64 TAGLFPBLAQVTFQEEAMSSPEVIFVAVFEHYSSLSGLADQLAGKILVDVSNPTEKRL 123
Qy 124 QYPSNAYLASLPDLSLIVGFNVVSAWALQGPXASROYICSNNTQARQOYIELAR 183
Db 124 QYPSNAYLASLPDLSLIVGFNVVSAWALQGPXASROYICSNNTQARQOYIELAR 183
Qy 124 QYPSNAYLASLPDLSLIVGFNVVSAWALQGPXASROYICSNNTQARQOYIELAR 183
Db 124 QYPSNAYLASLPDLSLIVGFNVVSAWALQGPXASROYICSNNTQARQOYIELAR 183
Qy 184 QINFTPIDLSLSAREIENPLRLFTLMRGPVVAISLATFFLYSPVVDVHYPRANQ 243
Db 184 QINFTPIDLSLSAREIENPLRLFTLMRGPVVAISLATFFLYSPVVDVHYPRANQ 243
Qy 184 AMGSTPILDMGSLASAREVEALPLRLFSMVKVPTLLMGLFVCFTYTNFTRDVLQPIRKO 243
Db 244 QSDPYKPIETVNTKLPVAVATLISLVYLAGLAAVQVLYGKXRRPFWLSTWLOCRK 303
Qy 244 ENKFKYKPLSVNTTLPVAVATLISLVYLAGLAAVQVLYGKXRRPFWLSTWLOCRK 303
Db 244 ENKFKYKPLSVNTTLPVAVATLISLVYLAGLAAVQVLYGKXRRPFWLSTWLOCRK 303
Qy 304 QLGILSFFFAVHVAVSLCLPFRSERLFLNMAVQCHANIENANNEEYTRLEMYISF 363
Db 304 QLGILSFFFAVHVAVSLCLPFRSERLFLNMAVQCHANIENANNEEYTRLEMYISF 363
Qy 364 QIGLISFFFAVHVAVSLCLPFRSERLFLNMAVQCHANIENANNEEYTRLEMYISF 423
Db 364 QIGLISFFFAVHVAVSLCLPFRSERLFLNMAVQCHANIENANNEEYTRLEMYISF 423

```



OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Serru V., Lambdin D., Lenoir C., Manivet P., Vaudourolle M.,  
 RA Kellermann O., Loric S.;  
 RT "Molecular cloning and expression of mouse dudulin 2,"  
 RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AY029586; AAK50539.1; -  
 DR MGD; MGI:1915678; Taap6.  
 DR GO; GO:0006118; P:electron transport; IEA.  
 DR InterPro; IPR004455; NADPoxred\_F420.  
 DR Pfam; PF03807; F420\_oxidored; 1.  
 SQ SEQUENCE 514 AA; 57268 MW; 33986C288AEC0E2 CRC64;

Query Match 55.6%; Score 1306; DB 11; Length 514;  
 Best Local Similarity 56.2%; Pred. No. 8.3e-97;  
 Matches 248; Conservative 81; Mismatches 100; Indels 12; Gaps 4;

QY 6 MMGSPKSLSETCLENGINGIDARKVTYVIGSGDFAKSLITRLIRGYNHVIIGSRNPKF 65  
 DB 14 LVSDDSGLAE--VP-----KEAPK--VGILSGDFARSLATRLVSGSFVWVGSRRPKR 63  
 QY 66 ASEFPFHVVDVTHEDALTKTNIIFVAIHREHYTSLMDLRLHLLVKGILIVSNMRIN-- 123  
 DB 64 TAGLFPSLAQVTFQBEAVSSPEVIFVAVFREHYSSLCISLDQLAGKILVDSNPTEKHL 123  
 QY 124 QYPSNAEYLALEPDSILVGFNVVSAALQLGPKDASROYICSNNIQARQVIELAR 183  
 DB 124 QHRQSNAYLASLPACTVYKAFNVISAMLLQAVPRGNQVILCGNDSKAKRWEMAR 183  
 QY 184 QLANFIPIDIGLSAREIENPLRLFTLMRGPVVAISLATFPFLYSFVADVHPHARNO 243  
 DB 184 AMGTPTDMGSLASAREVALPRLPLPSGKVPITLALGLFVCFYTVFIDVLCPIYRKD 243  
 QY 244 QSDPYKIPLEIVNKTLPYVATITLSLVYAGLAAAYQLYGYFRFPMLEMTQCRK 303  
 DB 244 ENKRYKPLSVNTTLQVAVYLLSLVLPGLVLAALQLRGITKQFPMLDHMLQHRK 303  
 QY 304 QIGLSFFPFAHVAISLCPMRSEERYFLNMAVQVHANINENNEEVRIMYISF 363  
 DB 304 QIGLSFFPFAHVAISLCPMRSEERYFLNMAVQVHANINENNEEVRIMYISF 363  
 QY 364 GIMSGLISLAVTISVSNAIWMREFSFOSTLGYVALLISTHVLITGMKRAEEY 423  
 DB 364 GVLMGLSLAVTISVSNAIWMKEFSFOSTLGYVALLISTHVLITGMKRAEEY 423  
 QY 424 YRFYTPNPFVLAIVLSIVIL 444  
 DB 424 YKFLPPTFTLTLLPCVIL 444

## RESULT 9

Q86SF6 PRELIMINARY; PRT; 488 AA.  
 AC Q86SF6;  
 DT 01-JUN-2003 (TREMBlrel. 24, Created)  
 DT 01-JUN-2003 (TREMBlrel. 24, Last sequence update)  
 DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)  
 DE Similar to dudulin 2 (TSA6).  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Kidney;  
 RA Straubeberg R.;  
 RL Submitted (JAN-2003) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=22506415; PubMed=12606722;  
 RA Passet B.V., Nancy-Porebois V., Amallag N., Prieur S., Cans C.,  
 RA Roborel de Climens A., Finucci G., Bouvard V., Tuvrder M., Susini L.,

RA Morchoiane S.P., Criddle V., Leespagnol A., Dausset J., Oren M.,  
 RA Amson R., Telerman A.;  
 RT "The p53-inducible TSAP6 gene product regulates apoptosis and the cell  
 cycle and interacts with Nix and the Myt1 Kinase,"  
 RL Proc. Natl. Acad. Sci. U.S.A. 100:2284-2289(2003).  
 DR EMBL; BC042150; AAH42150.1; -  
 DR EMBL; AY214461; AAO38238.1; -  
 DR GO; GO:0006118; P:electron transport; IEA.  
 DR InterPro; IPR03006; IG\_MHC.  
 DR InterPro; IPR004455; NADPoxred\_F420.  
 DR Pfam; PF03807; F420\_oxidored; 1.  
 DR PROSITE; PS00290; IG\_MHC; 1.  
 SQ SEQUENCE 488 AA; 54600 MW; C89EBD0D403F9FB CRC64;

Query Match 54.1%; Score 1272; DB 4; Length 488;  
 Best Local Similarity 54.5%; Pred. No. 4.3e-94;  
 Matches 242; Conservative 82; Mismatches 108; Indels 12; Gaps 4;

QY 3 SISMGSPKSLSETCLENGINGIDARKVTYVIGSGDFAKSLITRLIRGYNHVIIGSRN 62  
 DB 11 SLHVDSDSLAK--VPD-----EAPK--VGILSGDFARSLATRLVSGSFVWVGSRRPKR 60  
 QY 63 KPFASEFPFHVVDVTHEDALTKTNIIFVAIHREHYTSLMDLRLHLLVKGILIVSN--NM 120  
 DB 61 PRTARLPESAQVTFQBEAVSSPEVIFVAVFREHYSSLCISLDQLAGKILVDSNPTEQ 120  
 QY 121 RINQYPSNAEYLALEPDSILVGFNVVSAALQLGPKDASROYICSNNIQARQVIELAR 180  
 DB 121 EHLQHRQSNAYLASLPACTVYKAFNVISAMLLQAVPRGNQVILCGNDSKAKRWEMAR 180  
 QY 181 IARQNFIPIDIGLSAREIENPLRLFTLMRGPVVAISLATFPFLYSFVADVHPHARNO 240  
 DB 181 MALANGFPMVDGSLASAREVAMPRLPLPAKVPTLALGLFVCFYTVFIDVLCPIYRKD 240  
 QY 241 RNOQSDPYKIPLEIVNKTLPYVATITLSLVYAGLAAAYQLYGYFRFPMLEMTQCRK 300  
 DB 241 QESQNKFFPLPVSVNTTLQVAVYLLSLVLPGLVLAALQLRGITKQFPMLDHMLQHRK 300  
 QY 301 CRKQGLISLAVTISVSNAIWMREFSFOSTLGYVALLISTHVLITGMKRAEEY 360  
 DB 301 HKQIGLSLFFPFAHVAISLCPMRSEERYFLNMAVQVHANINENNEEVRIMYISF 360  
 QY 361 ISFGMSGLISLAVTISVSNAIWMREFSFOSTLGYVALLISTHVLITGMKRAEEY 420  
 DB 361 ISLGYLAAGTSLAVTISVSNAIWMKEFSFOSTLGYVALLISTHVLITGMKRAEEY 420  
 QY 421 EYRFYTPNPFVLAIVLSIVIL 444  
 DB 421 ESRYKFLPPTFTLTLLPCVIL 444

## RESULT 10

Q72389 PRELIMINARY; PRT; 498 AA.  
 AC Q72389;  
 DT 01-OCT-2003 (TREMBlrel. 25, Created)  
 DT 01-OCT-2003 (TREMBlrel. 25, Last sequence update)  
 DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)  
 DE Hypothetical protein DKFZP686H07150.  
 GN DKFZP686H07150.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Human colon endochei primary cell culture;  
 RA Bloeker H., Boecher M., Mewes H.W., Weill B., Amid C., Oeanger A.,  
 RA Robo G., Han M., Wiemann S.;  
 RL Submitted (JUN-2003) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; BX538047; CAD97986.1; -  
 RX Hypothetical protein.  
 SQ SEQUENCE 498 AA; 55593 MW; 69DD23D5A00C2DEB CRC64;

```

Query Match Similarity      54.1%; Score 1272; DB 4; Length 498;
Best Local Similarity      54.5%; Pred. No. 4, 4e-94;
Matches 242; Conservative  82; Mismatches 108; Indels 12; Gaps 4

QY 3 SISMGSPKSLSETCLPENGINGIKDKARKVTAVIGSGDFPAKSLTTRLRCGYHVGISRN 62
DB 21 SLHLVDSOSSLAK--VPD-----BAFK--VGILSGDFARSLATRLVSGGKVVGSRN 70
QY 63 PKFAEFPFHVDVTHHEHDLTKTNIPFAHREHYTSLMDRLHLYGKILVDYEN--NM 120
DB 71 PKRTARLPFSAQVTPQEBAVSPEVIFAVAREHYTSLICSLSDLAGKILVDYENPTDQ 130
QY 121 RINQYPSNAEYLIASLFPDLSLVKGFNVVSAVALQGPDKASROYVICSNNIOARQVIE 180
DB 121 EHLQHRSSNAEYLIASLFPCTCVKAFNVISAMTLAGPDRGNQVPICODGEAKRAVSE 190
QY 181 LARQALNFPIDLGSSASAREINMLRLFTLRGVVAVIASLATEPEFLSPFDVIHYA 240
DB 191 MALNAGFMEVDGSLASAEVEMAPKRLIPAKVFTTLALGLFVCFYAKNFRDVLQFV 250
QY 241 RNQGSDFPKPIPIETVNTKLPVIALFLLSLVYAGLIAAVALQYGNKYRRPPEMLTQL 300
DB 251 QESQKPKFKLPVSVVNTLLPCVAVYLLSLVLPVTLAALQRRGKTKRPFDDMLDHLQ 310
QY 301 CRKQGLISFPFAMHVAVSLCLPMRPSRRYFLMMAVQVHANENSMNEEVRIRIEM 360
DB 311 HRKQGLISFPFALAHVLSFCLPRRARVYLVMLAVQVLANSHLMVEEVRIRIEM 370
QY 361 ISFGISGLISLILAVTISPVSNALNMRRESFISITGLYVALLISTFHVILYMKRAFE 420
DB 371 LSLGLVALGTLISLAVTSLPSTANSLNMRRESFVQSLGFAVALVSTLHTLTVGNTRAFE 430
QY 421 ERYRFPYPPNVLAVLPSTIYL 444
DB 431 ESRYKFXLPPTFTLTLVPCVIL 454

RESULT 11
Q9NVB5 PRELIMINARY; PRT; 488 AA.
ID Q9NVB5
AC Q9NVB5;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Hypoetical protein FL110829 (Dudulin 2).
OS Homo sapiens (Human)
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Isegai T., Ota T., Hayashi K., Sugiyama T., Otsuki T., Suzuki Y.,
RA Nishikawa T., Nagai K., Sugano S., Aotsuka S., Yoshikawa Y.,
RA Marunawa H., Ishii S., Kawai Y., Saito K., Yamamoto U., Wakamatsu A.,
RA Nakamura Y., Nagahari K., Masuho Y., Sasaki N.,
RA "NEO human cDNA sequencing project."
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Serru V., Manivet P., Lenoit C., Eschwege P., Lambdin D.,
RA Vaudouin J., Kellermann O., Loric S.;
RT "Dudulin 2, a new tumor antigen expressed in various human tumors.",
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL, AK001691; BAA91839.1; -
DR EMBL, AY029585; AAK50538.1; -
DR GO; GO:0006118; P:electron transport; IEA.
DR InterPro; IPR004306; IG_MHC.
DR InterPro; IPR004455; NADPoxred_F420.
DR Pfam; PF03807; F420 oxidored_1.
DR PROSITE; PS00290; IG_MHC_1.
DR Hypoetical protein.
SQ SEQUENCE 488 AA; 54616 MW; EC0BCA483335AAD6 CRC64;

```

Query Match	54.0%	Score 1269	DB 4	Length 488
Best Local Similarity	54.3%	Pred. No. 7.5e-94		
Matches 241	Conservative 83	Mismatches 108	Indels 12	Gaps 4
QY	3	SISMMGSPKSLSECTPNCINGINGIKDARKYGVYIGSGDFAKSLTIRLRICGYHVIIGSRN	62	
DB	11	SLHLVDSDSLAK--VPD-----EAPK--VGIIIGSGDFASLSTRIVGSGFKVVGSSRN	60	
QY	63	PKRASEFPFHVVVDVTHHEBALTKNTIIIVAIHREHYTSSIMDLRLHVLGKILLIVSN--NM	120	
DB	61	PKRTAALYPSAAQVTFQEBANSSPVYIVAAVREHYTSSLSLSPQLGKILLIVDSNPTQ	120	
QY	121	RINQYEPESNAEYIASLFPDPSILVYGFNVYSAMALQGEKASRQYVICSNNIGARQVIE	180	
DB	121	EHQHRESNAEYIASLFPCTCVKAFNYIASMTLQAGPRDNRQVPCIGQPEAKRAVSE	180	
QY	181	LAQLNLFIPIIDGSLSSAREINLRLFTLMRGVVVAISLAFFFLYSFVVDVHPVA	240	
DB	181	MLAMGMFMDVDSGLASAEVEMPLRLPLMKVFTLALCLFVCFAYINVRDLQPV	240	
QY	241	RNOQSPFYKPIPIEIVNKTPIYATILSLVYLAGLAAAYOLYGYTKRRFPFWLTLQ	300	
DB	241	QSQNPFPLPYSVVNTLPCAAVYLLSVLPGVLAALQLRGTKXQFPPDWLDMWLQ	300	
QY	301	CRQQLGLSFPFAMVAVASLCPKRSEERTLFLNMAVQVHNAINENNEEYWRLEMV	360	
DB	301	HRQIIGLSFPCALALVSPCLPRARRARDVNLAKQVLAKSHLWEEBEVWEIY	360	
QY	361	ISFGMISGLSLHNTSIPSVNALNMRREFSFGTSTGYVALLISTFHVLYIGMKAFE	420	
DB	361	LSLGVIALGLSLHNTSIPSLANLNMRREFSFGVSSGFALVLSLHTLITVGTWRAFE	420	
QY	421	EERYRPTPPNFVLAIVPSYIYL 444		
DB	421	ESRYKRYLPPTFTLLVPCVIL 444		
RESULT 12				
Q8NEW6				
ID	Q8NEW6	PRELIMINARY;	PRT;	487 AA.
AC	Q8NEW6;			
DT	01-OCT-2002 (TrEMBLrel. 22, Created)			
DT	01-OCT-2002 (TrEMBLrel. 22, Last sequence update)			
DT	01-JUN-2003 (TrEMBLrel. 24, Last annotation update)			
DE	Phyde.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.			
NC	NCBI_TaxID=9606;			
NP	[1]			
RA	SEQUENCE FROM N.A.			
RA	Lu Y., Beheasti B., Squire J.A., Yang X.J.;			
RT	"Characterization of a novel apoptosis-inducing gene, hpnHyde, that			
RT	inhibits prostate cancer cell growth";			
RL	Submitted (MAR-2002) to the EMBL/Genbank/DBJ databases.			
DR	EMBL; AF082673; AA045136.1; ..			
DR	GO; GO:0006118; P:electron transport; IDA.			
DR	InterPro; IPR003006; IG MHC.			
DR	InterPro; IPR004455; NADPoxred.F420.			
DR	Pfam; PF03807; F420 oxidored; I.			
DR	PROSITE; PS00290; IG MHC; 1.			
SO	SEQUENCE 487 AA; 54501 MW; E0F95BE855C81EDF CRC64;			
Query Match	53.3%	Score 1252.5	DB 4	Length 487
Best Local Similarity	54.1%	Pred. No. 1.6e-92		
Matches 240	Conservative 83	Mismatches 108	Indels 13	Gaps 5
QY	3	SISMMGSPKSLSECTPNCINGINGIKDARKYGVYIGSGDFAKSLTIRLRICGYHVIIGSRN	62	
DB	11	SLHLVDSDSLAK--VPD-----EAPKVS--ILGSGDFASLSTRIVGSGFKVVGSSRY	60	
QY	63	PKRASEFPFHVVVDVTHHEBALTKNTIIIVAIHREHYTSSIMDLRLHVLGKILLIVSN--NM	120	

```
Db 61 PRTARLPFSAQVTFQEEAVSSPEVITFAVREHYSSLSISDQAGKILVDVSNPTQ 120
Qy 121 RINQYPSNAEYLAIFPDSLIVKGFNVSAVALQGPDAARQYVICSNNIOARQYIE 180
Db 121 EHLQRESNAEYLAIFPCTVYKAFNVISAWTLQAGPRDNGQVPCIGDQPAKRAVSE 180
Qy 181 IARQINFPIDIGSSASAREINLPLRLFTLMRGVVAISLATFFLYSTRDYIHYA 240
Db 181 MALAMGFVMDGSLASAVEVAMPRLPLPAKVPTLLALGFCFYANFVRDVLQPYV 240
Qy 241 RNOQSDPYKIPLEIVNKTLPVATITLSIVYLAGLAAAYOLYGYTKRFRPMTETIQ 300
Db 241 QESQNKFFKLPVSVNTTLPVAYVLLSVYLPGLAALQLRGTQKQRFPMDLHMQLQ 300
Qy 301 CRKQGLISFFPAMVAVASLCLPMRSERYFLNMAVOQVHANINSNWEEVRIEM 360
Db 301 HRKQIGLSFFCAALHAYSFCLPLRARHYDVLNLAAYQVLANKSHLM-VEEYWRMEIY 359
Qy 361 ISFGMSLGLSLAVTSIPSVNALNMBRFSFGTGLYVALLSTFPHVLYGMKRAFE 420
Db 360 ISLGVIALGTLISLATVSLPSIANSINMBRFSFVQSLGFVALVSTLHTLTGYWTRAFE 419
Qy 421 EBYRFPYTPNFVALVPSIVIL 444
Db 420 ESRYKFLPPTFTLTLLVPCVIL 443
```

## RESULT 13

```
Q8TF03 PRELIMINARY; PRT; 487 AA.
AC Q8TF03;
DT 01-JUN-2002 (Tremblrel. 21, Created)
DT 01-JUN-2002 (Tremblrel. 21, Last sequence update)
DT 01-JUN-2003 (Tremblrel. 24, Last annotation update)
DE Tumor suppressor pHyde.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Steiner M.S., Allay J.A., Wang C.;
RT "A novel prostate-derived tumor suppressor."
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF238864; AAL78206.1; -.
DR GO; GO:0006118; P:electron transport; IEA.
DR InterPro; IPR003006; IG_MHC.
DR InterPro; IPR004455; NADPoxred_F420.
DR Pfam; PF03807; F420 oxidored; I.
DR PROSITE; PS00290; IG_MHC; 1.
SQ SEQUENCE 487 AA; 54402 MW; AFP16053590E6F68 CRC64;
```

Query Match 53.0%; Score 1245.5; DB 4; Length 487;

Best Local Similarity 53.8%; Pred. No. 5.8e-92; Matches 239; Conservative 83; Mismatches 109; Indels 13; Gaps 5;

```
Qy 3 SISMGSPKSLSETCLPNGINGIKDARKVTGVISSGDFAKSLTIRLCGYHVVIGSRN 62
Db 11 SLHLVDSDSLAK--VPD-----EAPKVS--ILSGDFARSLATRLVSGSFKVVGSRN 60
Qy 63 PKFASEFPFHVVDTTHEDALTKNIIIFVAIHREHYTSLMDLRHLVKGKILVDVSN--NM 120
Db 61 PRTARLPFSAQVTFQEEAVSSPEVITFAVREHYSSLSISDQAGKILVDVSNPTQ 120
Qy 121 RINQYPSNAEYLAIFPDSLIVKGFNVSAVALQGPDAARQYVICSNNIOARQYIE 180
Db 121 EHLQRESNAEYLAIFPCTVYKAFNVISAWTLQAGPRDNGQVPCIGDQPAKRAVSE 180
Qy 181 IARQINFPIDIGSSASAREINLPLRLFTLMRGVVAISLATFFLYSTRDYIHYA 240
Db 181 MALAMGFVMDGSLASAVEVAMPRLPLPAKVPTLLALGFCFYANFVRDVLQPYV 240
```

```
Qy 241 RNOQSDPYKIPLEIVNKTLPVATITLSIVYLAGLAAAYOLYGYTKRFRPMTETIQ 300
Db 241 QESQNKFFKLPVSVNTTLPVAYVLLSVYLPGLAALQLRGTQKQRFPMDLHMQLQ 300
Qy 301 CRKQGLISFFPAMVAVASLCLPMRSERYFLNMAVOQVHANINSNWEEVRIEM 360
Db 301 HRKQIGLSFFCAALHAYSFCLPLRARHYDVLNLAAYQVLANKSHLM-VEEYWRMEIY 359
Qy 361 ISFGMSLGLSLAVTSIPSVNALNMBRFSFGTGLYVALLSTFPHVLYGMKRAFE 420
Db 360 ISLGVIALGTLISLATVSLPSIANSINMBRFSFVQSLGFVALVSTLHTLTGYWTRAFE 419
Qy 421 EBYRFPYTPNFVALVPSIVIL 444
Db 420 ESRYKFLPPTFTLTLLVPCVIL 443
```

## RESULT 14

```
Q8TDP3 PRELIMINARY; PRT; 456 AA.
AC Q8TDP3;
DT 01-JUN-2002 (Tremblrel. 21, Created)
DT 01-JUN-2002 (Tremblrel. 21, Last sequence update)
DT 01-JUN-2003 (Tremblrel. 24, Last annotation update)
DE pHyde II.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Wang C., Allay J.A., Steiner M.S.;
RT "Second human member of pHyde family, Human pHyde II."
RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF262322; AAM08128.1; -.
DR GO; GO:0006118; P:electron transport; IEA.
DR InterPro; IPR003006; IG_MHC.
DR InterPro; IPR004455; NADPoxred_F420.
DR Pfam; PF03807; F420 oxidored; I.
DR PROSITE; PS00290; IG_MHC; 1.
SQ SEQUENCE 456 AA; 50430 MW; CSF7C7008D55251E CRC64;
```

Query Match 45.7%; Score 1074.5; DB 4; Length 456;

Best Local Similarity 52.9%; Pred. No. 3.2e-78; Matches 209; Conservative 75; Mismatches 98; Indels 13; Gaps 5;

```
Qy 3 SISMGSPKSLSETCLPNGINGIKDARKVTGVISSGDFAKSLTIRLCGYHVVIGSRN 62
Db 11 SLHLVDSDSLAK--VPD-----EAPKVS--ILSGDFARSLATRLVSGSFKVVGSRN 60
Qy 63 PKFASEFPFHVVDTTHEDALTKNIIIFVAIHREHYTSLMDLRHLVKGKILVDVSN--NM 120
Db 61 PRTARLPFSAQVTFQEEAVSSPEVITFAVREHYSSLSISDQAGKILVDVSNPTQ 120
Qy 121 RINQYPSNAEYLAIFPDSLIVKGFNVSAVALQGPDAARQYVICSNNIOARQYIE 180
Db 121 EHLQRESNAEYLAIFPCTVYKAFNVISAWTLQAGPRDNGQVPCIGDQPAKRAVSE 180
Qy 181 IARQINFPIDIGSSASAREINLPLRLFTLMRGVVAISLATFFLYSTRDYIHYA 240
Db 181 MALAMGFVMDGSLASAVEVAMPRLPLPAKVPTLLALGFCFYANFVRDVLQPYV 240
Qy 241 RNOQSDPYKIPLEIVNKTLPVATITLSIVYLAGLAAAYOLYGYTKRFRPMTETIQ 300
Db 241 QESQNKFFKLPVSVNTTLPVAYVLLSVYLPGLAALQLRGTQKQRFPMDLHMQLQ 300
Qy 301 CRKQGLISFFPAMVAVASLCLPMRSERYFLNMAVOQVHANINSNWEEVRIEM 360
Db 301 HRKQIGLSFFCAALHAYSFCLPLRARHYDVLNLAAYQVLANKSHLM-VEEYWRMEIY 359
Qy 361 ISFGMSLGLSLAVTSIPSVNALNMBRFSFGTGLYVALLSTFPHVLYGMKRAFE 420
Db 360 ISLGVIALGTLISLATVSLPSIANSINMBRFSFVQ 394
```

## RESULT 15

Q923B6

ID Q923B6 PRELIMINARY; PRT; 470 AA.

AC Q923B6;

DT 01-DEC-2001 (TREMBLrel. 19, Created)

DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)

DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)

DE Similar to hypothetical protein FJ23153 (Tnfr-induced adipose-related

protein).

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OX NCBI\_TaxID=10090;

RN [1]

RP SEQUENCE FROM N.A.

RA Strausberg R.;

RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.

RN [2]

RP SEQUENCE FROM N.A.

RA STRAIN=C57BL/6J; TISSUE=Morris, and Vein;

RX MEDLINE=22354683; PubMed=12466851;

RA The FANTOM Consortium,

RA the RIKEN Genome Exploration Research Group Phase I &amp; II Team;

RT "Analysis of the mouse transcriptome based on functional annotation of

60,770 full-length cDNAs."

RL Nature 420:563-573 (2002).

DR EMBL: BC006651; AA006651.1; -

DR EMBL: AK040760; BAC30696.1; -

DR GO: 0006118; Electon transport; IEA.

DR InterPro: IPR000585; Hemopexin.

DR InterPro: IPR004455; NADPoxred; P420.

DR Pfam: PF03807; F420 oxidored; I.

DR PROSITE: PS00024; HEMOPEXIN; 1.

SQ SEQUENCE 470 AA; 52994 MW; 6823E7682AC78B19 CRC64;

Query Match 45.3%; Score 1064.5; DB 11; Length 470;

Best Local Similarity 48.5%; Pred. No. 2.1e-77;

Matches 200; Conservative 80; Mismatches 131; Indels 1; Gaps 1;

```
QY 33 VGVIGSGDPFASLTIRLIRGQHVIVIGSRNPKASEFPFHVVDVTHEDALTKTNIIPVA 92
DB 22 VCIFGTGDKSLGLMLDQGYISIVFGSRNPQ-VSLP-RGAEVLSYSEAAKSDIITLA 80
QY 93 IREHYTSLMDLRHLVKGKILIDVSNMKNINQYPSNAEYIASLPDGLIVGFNVVSAM 152
DB 81 MEREHYDSLTELVDYLKKGKVLVDVSNMKNINQYPSNAEYIASLPDGLIVGFNVVSAM 140
QY 153 ALOGPKASROYVICSNNICARQOYIELARQANFIPIDGLSSARREINPLRFTLM 212
DB 141 ALQSGTLDASROYFVCGNDKAKORWDIARTLGLPLDQSLMAASSIENYPLQLFPMW 200
QY 213 RGVVVAISLATEFFFLYSFVRDVIHPYARNQSDFFYKIPIEIVNKTLPYVAITLLSLYL 272
DB 201 RPFVYLSVLCVFFFYCAIREVLYPVNGKTDATYRLAISIPNRVFPITLILALAYL 260
QY 273 AGILAAAYOLYGTGKRRRPPWLETLQCRKQGLISFFPAMVAVYSLCLPMRSERYL 332
DB 261 PGLIAAILQLYRGKTRRRPPNWLDPHMLCRKQGLVAFGLAVLYTLVPIRIYRW 320
QY 333 FLNMAVQVHANINENSWNEEYWRLEMYISFGIMSIGLISLAVTISPSVSNALNMRPS 392
DB 321 LKNATITQALTNKDSPTISYAMINDSYALIGLGFLLIGLITSLPSVSNMVMWRFR 380
QY 393 FIOSTLGYVALLISTFHVLIYGMKRAFEERYRFTYPPNFVATLVPSIVIL 444
DB 381 FVQSKGLYTLVCTAHTLVYGGKRLPSILRMSLPSAYIATLVIPCAVIV 432
```

Search completed: March 1, 2004, 23:34:42  
Job time : 89 secs